*C h a p t e r 2*

# CHARACTERIZATION OF MICROBIAL ASSOCIATIONS WITH METHANOTROPHIC ARCHAEA AND SULFATE-REDUCING BACTERIA THROUGH STATISTICAL COMPARISON OF NESTED MAGNETO-FISH ENRICHMENTS

## Abstract

Methane seep systems along continental margins host diverse and dynamic microbial assemblages, sustained in large part through the microbially mediated process of sulfate-coupled Anaerobic Oxidation of Methane (AOM). This methanotrophic metabolism has been linked to a consortia of anaerobic methane-oxidizing archaea (ANME) and sulfate-reducing bacteria (SRB). These two groups are the focus of numerous studies; however, less is known about the wide diversity of other seep associated microorganisms. We selected a hierarchical set of FISH probes targeting a range of *Deltaproteobacteria* diversity. Using the Magneto-FISH enrichment technique, we then magnetically captured CARD-FISH hybridized cells and their physically associated microorganisms from a methane seep sediment incubation. DNA from nested Magneto-FISH experiments was analyzed using Illumina tag 16S rRNA gene sequencing (iTag). Enrichment success and potential bias with iTag was evaluated in the context of full-length 16S rRNA gene clone libraries, CARD-FISH, functional gene clone libraries, and iTag mock communities. We determined commonly used Earth Microbiome Project (EMP) iTAG primers introduced bias in some common methane seep microbial taxa that reduced the ability to directly compare OTU relative abundances within a sample, but comparison of relative abundances between samples (in nearly all cases) and whole community-based analyses were robust. The iTag dataset was subjected to statistical co-occurrence measures of the most abundant OTUs to determine which taxa in this dataset were most correlated across all samples. Many non-canonical microbial partnerships were statistically significant in our co-occurrence network analysis, most of which were not recovered with conventional clone library sequencing, demonstrating the utility of combining Magneto-FISH and iTag sequencing methods for hypothesis generation of associations within complex microbial communities. Network analysis pointed to many co-occurrences containing putatively heterotrophic, candidate phyla such as OD1, Atribacteria, MBG-B, and Hyd24-12 and the potential for complex sulfur cycling involving *Epsilon-*, *Delta-*, and *Gammaproteobacteria* in methane seep ecosystems.

## Introduction

A central goal in microbial ecology is identifying and understanding microbial interactions in the environment. This goal can be addressed at many scales from statistical analyses of entire ecosystems (Barberán et al. 2012; Malfatti & Azam 2010; Ruff et al. 2015; Steele et al. 2011; Sunagawa et al. 2015) to high resolution image analysis of specific symbioses (Malfatti & Azam 2010; McGlynn et al. 2015; Orphan 2009; Orphan et al. 2001b; Wegener et al. 2015). Previous studies have shown that complex datasets can be distilled to determine primary ecosystem drivers, such as temperature, as main predictors of community variability (Sunagawa et al. 2015). In addition to correlating microbial patterns to environmental factors, interspecies interactions can be evaluated with methods such as co-occurrence analysis (Friedman & Alm 2012). Statistical significance of co-occurrence can be assessed at scales ranging from the entire genome to the operational taxonomic unit (OTU) (Barberán et al. 2012; Chaffron et al. 2010).

Many physical separation methods have been developed to partition complex microbial assemblages before analysis, including fluorescence-activated flow sorting (Amann et al. 1990; Yilmaz et al. 2010), optical trapping (Ashkin 1997), microfluidics (Melin & Quake 2007), and immunomagnetic beads (Pernthaler et al. 2008; Šafařík & Šafaříková 1999) that use characteristics of interest such as phylogenetic identity (Fluorescence In-Situ Hybridization; FISH) or activity (Berry et al. 2015; Hatzenpichler & Orphan 2015; Hatzenpichler et al. 2014; Kalyuzhnaya et al. 2008; Wegener et al. 2012).

Here we combine Magneto-FISH and Illumina Tag (iTag) sequencing utilizing the Earth Microbiome Project (EMP) universal primer set (Caporaso et al. 2012). The Magneto-FISH method was originally developed to enrich for and characterize multi-species microbial associations in environmental samples (Pernthaler et al. 2008). This method consists of a liquid CARD (CAtalyzed Reporter Deposition)-FISH reaction as a 16S rRNA gene identity-based selection mechanism followed by an immunomagnetic sediment matrix separation mechanism to target specific phylogenetic groups in conjunction with their physically associated microbial partners. By combining this method for phylogenetically targeted physical separation with high throughput amplicon sequencing, we can compare an array of associated microbial communities in

parallel, with replicates. This provides statistical power in deriving microbial associations from complex sediment community assemblages, and thereby improving hypothesis development.

Anaerobic methane-oxidizing (ANME) archaea and sulfate-reducing *Deltaproteobacteria* (SRB) are the predominant community members discussed in methane seep literature and form syntrophic partnerships in physical associations, termed "aggregates" or consortia (Boetius et al. 2000; Green-Saxena et al. 2014; Knittel et al. 2003; Orphan et al. 2001a; Schreiber et al. 2010). Since physical association appears to be an important element for consortia activity (McGlynn et al. 2015; Wegener et al. 2015), methods like Magneto-FISH are ideal for probing this system because target organisms are separated from the sediment matrix along with their physically associated partners. A hierarchical probe set was chosen targeting *Deltaproteobacteria* and their ANME partners to create nested Magneto-FISH enrichments from methane seep sediment incubations under methane headspace. This method allows us to examine potential physical associations between ANME and SRB taxa and other microorganisms using co-occurrence statistical methods applied to iTag sequences from nested Magneto-FISH enrichments.

ANME have been broadly divided into three separate groups, which can be further subdivided into ANME-1a, 1b, 2a, 2b, 2c, and 2d, and 3. ANME-1 archaea are a unique order-level lineage within the *Euryarchaeota,* between the *Methanomicrobiales* and the *Methanosarcinales*, known to associate with sulfate-reducing bacteria, but obligately associated lineages have yet to be defined. ANME-2 archaea, within the order *Methanosarcinales*, commonly form associations with *Desulfosarcina*/*Desulfococcus*-related (DSS) sulfate-reducing *Deltaproteobacteria* (Boetius et al. 2000; Orphan et al. 2001a; Schreiber et al. 2010). They have also been found in association with *Desulfobulbus*-related (DSB) *Deltaproteobacteria* in the same environments, where geochemical factors have been suggested as a possible explanation for partner differentiation (Green-Saxena et al. 2014). ANME-2a/b and ANME-2c both predominately associate with a subgroup of DSS, SEEP-SRB1 (Schreiber et al. 2010), but also form consortia with DSB (Green-Saxena et al. 2014; Pernthaler et al. 2008). ANME-3 has been found in association with *Desulfobulbus*-related *Deltaproteobacteria* (Niemann et al. 2006) and SEEP-SRB1 (Schreiber et al. 2010). These ANME groups have also been observed in the environment without bacterial partners (House et al. 2009; Orphan et al. 2002; Schreiber et al. 2010; Treude et al. 2007). In addition to ANME archaea, other uncultured archaeal lineages commonly recovered from methane seeps include Marine Benthic

Group-D (*Thermoplasmatales*), Deep Sea Archaeal Group / Marine Benthic Group-B (Ruff et al. 2015; Yanagawa et al. 2011), and sometimes methanogens (Orphan et al. 2001a; Ruff et al. 2015; Takano et al. 2013; Vigneron et al. 2015).

*Deltaproteobacteria* diversity beyond DSS and DSB has also been well described in methane seeps. In addition to SEEP-SRB1, Knittel et al. (2003) define three more *Deltaproteobacteria* clades within *Desulfobulbaceae* (SEEP-SRB2, 3 and 4). Green-Saxena et al. (2014) also described a *Desulfobulbaceae* affiliated seepDBB group in methane seep systems. Bacterial diversity surveys of methane seep habitats frequently report occurrence of other diverse *Proteobacteria* including sulfur oxidizers (*Gammaproteobacteria* and *Epsilonproteobacteria*) and putative heterotrophs (*Alphaproteobacteria* and *Betaproteobacteria*) (Pernthaler et al. 2008; Ravenschlag et al. 1999). Many other bacterial phyla have also been found such as *Firmicutes*, *Thermomicrobia*, *Bacteroidetes*, *Chlorobi*, *Nitrospira*, WS3, OD1, OP11, TM7, and WS6 (Schreiber et al. 2010); *Cytophaga* and *Flavobacteria* (Knittel et al. 2003); *Chloroflexi*, *Atribacteria* (previously Candidate Division JS1), CD12, WS1, OS-K, AC1, and *Planctomycetes* (Yanagawa et al. 2011); and *Acidobacteria* (Ravenschlag et al. 1999). Ruff et al. (2015) indentify *Methanomicrobia*, *Deltaproteobacteria*, Hyd24-12 and Atribacteria as the characteristic 'core' microbial taxa in methane seep ecosystems, as compared to *Gammaproteobacteria*, *Flavobacteria*, *Thermoplasmatales*, and MBG-B taxa that were found in high relative abundance in seeps and other marine ecosystems.

Despite the wealth of bacterial and archaeal diversity in methane seep sediments, little is known about potential associations with ANME/SRB, or associations that do not involve ANME or SRB. Our study utilizes the novel combination of targeted Magneto-FISH enrichment of specific microbial taxonomic groups and iTag sequencing to develop statistically supported co-occurrence microbial networks to address knowledge gaps in our understanding of methane seep microbial communities. Network analysis revealed many novel associations between methane seep *Proteobacteria* taxa and Candidate phyla. The significant co-occurrences observed suggest new avenues for future studies on microbial interactions involved in carbon and sulfur cycling in methane seep systems.

**Materials & Methods**.

*Sample collection and Magneto-FISH*

iTag Magneto-FISH enrichments were conducted using a large scale (1 L) incubation of methane seep sediment from Hydrate Ridge North (offshore Oregon, USA) collected in September 2011 at 44°40.02' N 125°6.00' W, from a water depth of 775 m using the ROV *JASON II* and the R/V *Atlantis*. Marine sediment was collected using a push core to sample a sulfide-oxidizing microbial mat adjacent to an actively bubbling methane vent. A sediment slurry from the upper 0–15 cm depth horizon of the push core was prepared with 1 volume $N_2$ sparged artificial seawater to 1 volume sediment, overpressurized with methane (3 bar) and incubated at 8$^{o}$C in a 1 L Pyrex bottle capped with a butyl rubber stopper until subsampling for Magneto-FISH.

In February 2015, incubation samples were immediately fixed in 0.5 ml sediment aliquots in 2% paraformaldehyde (PFA) for 3 hrs at 4$^{o}$C. The samples were washed in 50% phosphate-buffered saline (PBS): 50% EtOH, then 75% EtOH: 25% DI water, and resuspended in 2 volumes (1 ml) 100% ethanol. Samples were centrifuged at 1000 × *g* for 1 min between wash steps. After fixation, the Magneto-FISH method first described by Pernthaler et al. (2008) and further optimized by Schattenhofer and Wendeberg (2011) and Trembath-Reichert et al. (2013) was used. Briefly, a liquid CARD-FISH reaction was followed by immunomagnetic bead incubation coupled with anti-fluorecsein attaching magnetic beads to CARD-FISH hybridized aggregates. Samples were then held against magnets and the sediment matrix was washed away, retaining target cells and physically associated microbes in the magnetic portion, as described in Trembath-Reichert et al. 2013. Four previously published FISH probes were used targeting a range of *Deltaproteobacteria* and *Methanomicrobia* (Table 1). A subset of three 0.5 ml aliquots was also immediately frozen before fixation (unfixed bulk sediment), and another three aliquots were frozen after fixation (fixed bulk sediment) for bulk sediment comparison with Magneto-FISH enrichments. Sediment for MSMX-Eel_932 Magneto-FISH metabolic gene analysis was fixed and washed onboard in September 2011, as described above. See methods flow chart provided in Sup Figure 1.

Table 1: FISH probes and primers used in this study. References: (Akerman et al. 2013; Blazejak et al. 2006; Boetius et al. 2000; Caporaso et al. 2012; DeLong 1992; Lane 1991; Loy et al. 2002; Macalady et al. 2006; Manz et al. 1996; Manz et al. 1992; Manz et al. 1998; Neef et al. 1998; Schreiber et al. 2010; von Wintzingerode et al. 1999; Wagner et al. 1998)

| *Name* | *Sequence (5' -> 3')* | *Target* | *FA (%) / Annealing ($^{o}$C)* |
|---|---|---|---|
| **PROBES for Magneto-FISH & CARD-FISH** | | | |
| DSS_658 | TCCACTTCCCTCTCCCAT | Desulfosarcina/ Desulfococcus, Desulfofaba, Desulfofrigus | 50 |
| Delta_495a | AGTTAGCCGGTGCTTCCT | Most Deltaproteobacteria and most Gemmatimonadetes | 35 |
| Delta_495a-comp | AGTTAGCCGGTGCTTCTT | | 35 |
| Seep-1a_1441 | CCCCTTGCGGGTTGGTCC | Seep-SRB1a | 45 |
| MSMX-Eel_932 | AGCTCCACCCGTTGTAGT | All ANME groups | 35 |
| ANME-1_350 | AGTTTTCGCGCCTGATGC | ANME-1 | 40 |
| Epsi_404 | AAAKGYGTCATCCTCCA | Epsilonproteobacteria | 30 |
| Gam_42a | GCCTTCCCACATCGTTT | Gammaproteobacteria | 35 |
| Gam_42a comp (Bet42a) | GCCTTCCCACTTCGTTT | Betaproteobacteria | 35 |
| Pla_46 | GACTTGCATGCCTAATCC | Planctomycetes | 35 |
| Pla_886 | GCCTTGCGACCATACTCCC | Planctomycetes | 35 |
| CF_319A | TGGTCCGTGTCTCAGTAC | CFB (Cytophaga, Bacteriodales, Flavobacterium, Sphingobacterium) | 35 |
| CF_319B | TGGTCCGTATCTCAGTAC | CFB (mostly Cytophaga) | 35 |
| **PRIMERS for iTAG** | | | |
| 515F | GTGCCAGCMGCCGCGGTAA | V4 region universal 16S rRNA | 55 |
| 806R | GGACTACHVGGGTWTCTAAT | V4 region universal 16S rRNA | 55 |
| **PRIMERS for CLONE LIBRARIES** | | | |
| Bac27F | AGAGTTTGATYMTGGCTC | Bacterial 16S rRNA | 54 |
| U1492R | GGYTACCTTGTTACGACTT | Universal 16S rRNA | 54 |
| 10-30Fa | TCCGGTTGATCCTGCC | Archaeal 16S | 54 |
| Arc958R | YCCGGCGTTGAMTCCAATT | Archaeal 16S | 54 |
| DSR1F | ACSCACTGGAAGCACG | dsrAB | 61-48 |
| DSR4R | GTGTAGCAGTTACCGCA | dsrAB | 61-48 |
| APS_1F | TGGCAGATCATGATYMAYGG | APS reductase | 54 |
| APS_4R | GCGCCAACYGGRCCRTA | APS reductase | 54 |
| sox527F | TGGTWGGWCAYTGGGAATTTA | sulfate thiol esterase | 46 |
| sox1198R | AGAANGTATCTCKYTTATAAAG | sulfate thiol esterase | 46 |

*iTag Amplification*

For iTag sequencing, ten Magneto-FISH enrichments were performed in parallel using the FISH probes DSS_658 (triplicate), MSMX-Eel_932 (triplicate), SEEP-1a_1441 (duplicate), Delta_495a + Delta_495a competitor (duplicate). Magneto-FISH enrichments and bulk sediment samples were resuspended in 650 µl solution PM1 and transferred to silica tubes from the PowerMicrobiome RNA Isolation Kit (MoBio). This kit was chosen based on manufacturer recommendation for formalin-fixed sediment samples, with the added capability to co-elute RNA if desired. 6.5 µl of beta-mercaptethanol was added, and samples were mechanically lysed in a bead beater (FastPrepFP120, ThermoElectronCorp.) for 45 s at setting 5.5 and incubated at 65$^{o}$C for 3.5 hrs. The remaining steps in the PowerMicrobiome RNA Isolation Kit were followed according to manufacturer instructions (starting at step 5) without any DNase procedures, and eluting in a final volume of 60 µl ultrapure water. DNA extracts were quantified using a Qubit Flurometer and HS dsDNA kit (Invitrogen; Sup Table 1). All but one Magneto-FISH sample had DNA concentrations below detection (<0.5 ng/µl); however, all samples yielded PCR amplicons when viewed on a gel after initial pre-barcoding PCR (30 cycles).

iTag samples were prepared with Earth Microbiome Project (EMP) primers 515f and 806r (Caporaso et al. 2012). An initial amplification of 30 cycles with primers lacking the barcode, linker, pad, and adapter was performed for all samples, in duplicate. Duplicate PCR reactions were pooled and reconditioned for 5 cycles with barcoded primers, for a total of 35 cycles. A master mix of 2X Q5 Hot Start High Fidelity Master Mix (NEB) and 10 µM forward and reverse primers was prepared for a final volume of 15 µl per sample, with 1 µl DNA template. PCRs had an initial 2 min heating step at 98$^{o}$C, followed by cycles of 10 s 98$^{o}$C, 20 s 54$^{o}$C, and 20 s 72$^{o}$C, and finished with a final extension of 2 min at 72$^{o}$C. PCR negative controls, substituting ultrapure water for DNA template, were amplified for 40 cycles total. We note that these are not the official recommended reagents or PCR conditions from the EMP, but internal lab tests showed that for 6 out of 9 mock community taxa, recovered sequence relative abundances were more accurate when using Q5 polymerase rather than the recommended Hot Start MasterMix (5-prime). EMP primers were chosen for iTag for cross-comparison between studies, though there is known primer bias within this universal primer set (Parada et al. 2015) and sequencing reactions will always have some inherent variability.

*Mock Communities*

Four mock communities were prepared with a range of relative proportions of nine common methane seep taxa (Sup Table 2). Full-length 16S rRNA gene plasmids from each taxa listed were quantified by Qubit. Taking into account the plasmid's nucleotide composition and length in order to calculate its molecular weight, plasmids were quantitatively combined in known volumetric fractions to achieve a range of desired mock community compositions. These combined plasmid mixes were diluted to ~1 ng/µL and then prepared according to the same iTag methods as all other samples.

*iTag sequence processing*

We followed the *mothur* Standard Operating Procedure (SOP) for Illumina MiSeq sequencing of the 16S rRNA gene V4 region, accessed May 2015 and using methods described in Kozich et al. (2013) with UCHIME chimera checking (Edgar et al. 2011). A concatenated file of the *mothur* version of separate archaeal and bacterial SILVA 119 databases (Quast et al. 2013) was used for alignment and classification. Unfixed Bulk Sediment 1 only returned 8% of the average DNA concentration of the other two samples. (Sup Table 1). This sample was removed from statistical analyses because it fails to be a representative of the unfixed bulk sediment community baseline. The mock communities were processed following the "Assessing Error Rates" section of the *mothur* SOP to compute sequencing error rates and spurious OTU rates (Sup Table 4). Additional analysis demonstrating sequence processing did not selectively remove ANME-2c sequences and relative sequence abundances recovered with iTag sequencing of mock communities are provided in Sup Table 3 and Sup Table 2, respectively.

Using *R* version 3.1.3 (R Core Team 2015), an average number of sequences per OTU was calculated from unfixed bulk sediment samples (2 and 3). All OTUs with an average relative sequence abundance below 0.1% in the unfixed bulk sediment were identified and removed from all samples using *mothur*. 135 unique OTUs remained out of 25,354. We also verified that after the 0.1% cutoff was applied, no negative control contaminant OTUs remained. The top 20 OTUs amplified from the no template negative control were classified as, in order of sequence

abundance: Sphingomonas*; Planctomyces*; Escherichia-Shigella*; Staphylococcus; Roseomonas*; Pir4_lineage; Delftia*; Macrococcus; Myxococcales;0319-6G20;unclassified; Planctomyces; Enhydrobacter; Sphingobium*; Caenispirillum; Bacillus*; Pseudoxanthomonas*; Peptoniphilus; Lysobacter; Salinicoccus; Propionibacterium.* Reagent contaminant genera discussed in Salter et al. (2014) are denoted by (*). All samples (including mock community and negative controls) were submitted to the SRA under the accession SAMN03879962, BioSample: SAMN03879962, Sample name: PC47 (5133-5137) mixed slurry.

Gene libraries of the Magneto-FISH samples were prepared as in Trembath-Reichert et al. (2013) using the primers and annealing temperatures listed in Table 1 and TOPO TA Cloning Kit for Sequencing with pCR4-TOPO Vector and One Shot Top 10 chemically competent Escherichia coli (Life Technologies). All full-length 16S rRNA gene sequences were aligned by the SINA online aligner (v 1.2.11; Pruesse et al. 2012) and added using maximum parsimony to the SILVA 119 database (Quast et al. 2013) for classification. A taxonomy-based count table was prepared (sequences per taxa, per sample) and all taxa absent from the bulk sediment library were removed from Magneto-FISH enrichment libraries (for parity with iTag contaminant removal processing). Functional gene sequences were translated using the *EMBOSS* online translation tool (Li et al. 2015), then added to *ARB* (Ludwig et al. 2004) databases for phylogenetic placement and classification. Sequences were submitted to NCBI under the following accession numbers: AprA (KT280505 - KT280517), DsrA (KT280518 - KT280533), McrA (KT280534 - KT280581), Archaeal 16S rRNA gene (KT280582 - KT280632), Bacterial 16S rRNA gene (KT280633 - KT280909), SoxB (KT280910 - KT280928). Gene trees were computed with representative sequences using PhyML 3.0 (Guindon et al. 2010) online execution with defaults on the South of France Bioinformatics platform.

*Statistical Analysis*

Weighted UniFrac (Lozupone & Knight 2005), Metastats (White et al. 2009), and linear discriminant analysis (LDA) effect size (LEfSe) (Segata et al. 2011) analyses were computed in *mothur* as outlined in the *mothur* SOP. Co-occurrence statistical analyses were run using the table of 135 unique OTUs in the format of sequence counts of each OTU per sample. The program *SparCC* was used to determine significant correlations (Friedman & Alm 2012). This analysis was

run 100 times with default settings, except 10 iterations were used instead of 20. OTUs with *SparCC* correlations above an absolute value of 0.6 with p-values below 0.01 were considered significant. Resulting associations that occurred in at least 50 out of 100 network iterations are provided in Sup Table 5. *Cytoscape* (Shannon et al. 2003) was used to display associations in Figure 1.

*CARD-FISH microscopy*

A triple CARD-FISH hybridization was performed with bacterial probes listed in Table 1, ANME-1_350 and MSMX-Eel_932. The sample preparation and CARD reaction was performed as per Green-Saxena et al. (2014). After the three CARD reactions, samples were post-stained with DAPI (25 ng/µl). CARD signal within any part of a physically attached group of cells larger than 10 µm was counted as a positive identification. For example, a large EPS matrix that contained many smaller separate ANME-1 and ANME-2 aggregates would count as one positive identification for each clade. This was done to simulate groups that would have been isolated together in a Magneto-FISH enrichment. Since the MSMX-Eel_932 probe also targets the ANME-1 population, only cells with MSMX-Eel_932 signal and no ANME-1_350 signal were recorded as an ANME-2 positive identification to comprehensively target ANME-1, -2, and a bacterial partner in a triple CARD-FISH hybridization set. ANME-3 were not recovered in the iTag dataset and were not considered as potential contributors to MSMX-Eel_932 signal.

**Results.**

*Relative sequence abundance of seep microbiome taxa in 16S rRNA gene iTag and libraries*

Relative sequence abundances of the methane seep microbiome characteristic taxa, ANME archaea, *Deltaproteobacteria*, Hyd24-12, and Atribacteria (Ruff et al. 2015), were compared two ways: 1) between iTag and gene library 16S rRNA gene samples to determine how relative sequence abundances differed between sequencing methodologies, and 2) between Magneto-FISH enrichment and bulk sediment to determine taxa-specific relative sequence abundance for each probe (Table 1).

Mock community analysis showed that ANME-2 were always underrepresented in iTag data (0.32-0.81 fold of what was expected), whereas the *Deltaproteobacteria* and ANME-1b were more faithfully represented (Sup Table 2). ANME-1a was consistently over amplified. By normalizing the relative sequence abundance of ANME-2c, -2a/b, and -1a to the abundance of ANME-1b, the most faithfully amplified archaea in the mock community data (Sup Table 2), we could compute a ratio between the average relative sequence abundance in fixed bulk sediment samples between iTag and the archaeal 16S rRNA gene library. ANME-2c (0.04 iTag:clone ratio), ANME-2a/b (0.12), and ANME-1a (0.40) were all less abundant in iTag sequences as compared to the archaeal gene clone library (calculated from values in Table 2). Similarly comparing SEEP-SRB1 to *Desulfobulbus* between the two methods in fixed bulk sediment returns a ratio of 0.41 iTag:clone. Since the iTag methodology recovers far more diversity (e.g. *Desulfobacula*, *Desulfocapsa*, *Desulfoluna*, Atribacteria, and Hyd24-12 were not recovered in the bacterial 16S rRNA gene bulk sediment library), it is expected that the relative sequence abundances of each individual taxon computed from iTag data would be less than from the domain targeted 16S rRNA gene libraries. However, the ANME-2c abundance ratio was an order of magnitude less than ANME-1a and SEEP-SRB1 ratios, and appears to be an extreme case of underestimation in iTag data. There was also variation between Magneto-FISH enrichment replicates, as indicated by the high standard deviations of Magneto-FISH samples as compared to bulk sediment samples. The degree of variation (average standard deviation across all taxa listed) correlated with the specificity of the probe; where Delta_495a had the lowest average standard deviation and Seep-1a_1441 had the highest average standard deviation.

The high relative sequence abundance taxa (>1.5 fold relative sequence abundance increase over fixed bulk sediment; Table 2) in the averaged Seep-1a_1441 iTag Magneto-FISH enrichments were *Desulfoluna* (2.20), SEEP-SRB1 (2.36), Hyd24-12 (3.44), and *Atribacteria* (1.51) (Table 2). The DSS_658 enrichment had fewer high relative sequence abundance taxa with only *Desulfoluna* (4.62), *Spirochaeta* (4.36), and *Atribacteria* (4.80). The Delta_495a enrichment also had three high relative sequence abundance taxa with *Desulfobulbus* (2.52), *Spirochaetae*-uncultured (3.70), and *Atribacteria* (3.02). The MSMX-Eel_932 enrichment had six high relative sequence abundance taxa with *Desulfococcus* (1.85), *Desulfoluna* (8.47), SEEP-SRB1 (1.67), *Spirochaeta* (1.63), Hyd24-12 (1.73), and *Atribacteria* (7.18). Gene library results showed high relative

sequence abundance (>1.5) in both ANME and *Deltaproteobacteria* with DSS_658 and MSMX-Eel_932 enrichments (Table 2). Similar to the bulk sediment, *Desulfobacula*, *Desulfocapsa*, *Desulfoluna*, *Atribacteria* and Hyd24-12 were not recovered in the bacterial 16S rRNA gene Magneto-FISH libraries. MSMX-Eel_932 enriched for SEEP-SRB1 (2.73), SEEP-SRB4 (3.28), *Desulfococcus* (3.82), *Spirochaeta* (1.64), and ANME-2a/b (2.51) in 16S rRNA gene libraries. There was also a slight enrichment of ANME-2c (1.28). The DSS_658 enrichment had high relative sequence abundance for SEEP-SRB1 (1.74), SEEP-SRB2 (2.78), ANME-2c (1.54), and ANME-2a/b (2.24) with iTag, but these same taxa did not have high relative sequence abundance in the gene library. *Spirochaeta* and SEEP-SRB1 had high relative sequence abundance in both iTag and gene libraries for MSMX-Eel_932 enrichments. Relative sequence abundances for all non-core methane seep taxa in iTag samples are included in Table 3, and where Magneto-FISH enrichments of these additional taxa support network co-occurrences they are discussed in network results.

*Statistical evaluation of Magneto-FISH enrichment*

To statistically compare enrichment microbial communities, we used a suite of statistical tests including: non-parametric T-tests (White et al. 2009), LEfSe (Segata et al. 2011), and UniFrac (Lozupone & Knight 2005). Using the T-test comparison, ten OTUs were significantly ($p<0.001$) different between the bulk sediment and Magneto-FISH samples (when only including OTUs with sequences present in both groups). The taxonomic assignments for these ten OTUs were: WCHB1-69, *Desulfobulbus*, *Thaumarcheota*, ANME-1a, *Bacteroidetes* (VC2.1), ANME-2c, *Caldithrix*, SEEP-SRB1, Candidate Division TA06, and *Gammaproteobacteria* (CS-B046). LEfSe was then used to determine which OTUs were significantly different between Magneto-FISH enrichments and bulk sediment. We found three OTUs were significantly (p-value <0.05) higher in relative sequence abundance in Magneto-FISH samples over bulk sediment with the taxonomies: SEEP-SRB1, *Desulfobulbus*, and *Planctomycetes* (SHA-43).

Weighted UniFrac analysis was used to compare the community composition between Magneto-FISH iTag enrichments. The UniFrac metric represents the fraction of the branch length that is unique to each sample, or unshared between samples, such that a higher ratio means less similar samples. The *Deltaproteobacteria* probe enrichment communities were more similar to each other

| | 16S rRNA gene (iTAG) | | | | | | | | | | | | | | | | 16S rRNA gene (Clone Library) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Seep1a_1441 | | | DSS_658 | | | Delta_495a | | | MSMX-Eel_932 | | | Fixed Bulk | | Unfixed Bulk | | DSS_658 | | MSMX-Eel_932 | | Fixed Bulk |
| Taxon | Avg. | Stdev. | Rel. Fixed | Avg. | Stdev. | Rel. Fixed | Avg. | Stdev. | Rel. Fixed | Avg. | Stdev. | Rel. Fixed | Avg. | Stdev. | Avg. | Stdev. | 24 arc, 41 bac | Rel. Fixed | 60 arc, 87 bac | Rel. Fixed | 43 arc, 95 bac |
| *ANME-1a* | 0.07 | 0.07 | 0.67 | 0.04 | 0.04 | 0.36 | | 0.01 | 0.05 | 0.07 | 0.01 | 0.61 | 0.11 | 0.02 | 0.10 | 0.02 | 0.08 | 0.28 | 0.08 | 0.28 | 0.30 |
| *ANME-1b* | 0.11 | 0.08 | 0.92 | 0.09 | 0.05 | 0.74 | 0.12 | 0.05 | 0.95 | 0.15 | 0.09 | 1.22 | 0.12 | 0.03 | 0.08 | 0.01 | | | | | 0.14 |
| *ANME-2a/b* | | 0.01 | 0.19 | 0.01 | 0.01 | 0.31 | | | 0.01 | | | 0.11 | 0.02 | 0.01 | 0.01 | | 0.42 | 2.24 | 0.47 | 2.51 | 0.19 |
| *ANME-2c* | | | 0.01 | | | 0.09 | | | | | | | 0.01 | 0.01 | 0.01 | | 0.50 | 1.54 | 0.42 | 1.28 | 0.33 |
| *Desulfobacula* | | | | | | | | | | | | | | | | | | | 0.01 | | |
| *Desulfobulbus* | 0.08 | 0.06 | 1.01 | 0.11 | 0.05 | 1.30 | 0.20 | 0.14 | 2.52 | 0.03 | 0.01 | 0.36 | 0.08 | 0.01 | 0.12 | 0.01 | 0.05 | 0.66 | 0.06 | 0.78 | 0.07 |
| *Desulfocapsa* | 0.02 | 0.01 | 1.02 | | | 0.16 | | 0.01 | 0.32 | 0.02 | 0.03 | 1.10 | 0.01 | | 0.01 | | 0.05 | | | | |
| *Desulfococcus* | 0.03 | 0.03 | 0.67 | 0.03 | 0.03 | 0.61 | 0.03 | 0.04 | 0.74 | 0.08 | 0.13 | 1.85 | 0.04 | | 0.03 | | | | 0.08 | 3.82 | 0.02 |
| *Desulfoluna* | 0.01 | 0.02 | 2.20 | 0.02 | 0.02 | 4.62 | | | | 0.04 | 0.04 | 8.47 | 0.01 | | 0.01 | | | | | | |
| *SEEP-SRB1* | 0.13 | 0.07 | 2.36 | 0.05 | 0.01 | 0.84 | 0.04 | 0.01 | 0.78 | 0.09 | 0.08 | 1.67 | 0.06 | | 0.06 | | 0.22 | 1.74 | 0.34 | 2.73 | 0.13 |
| *SEEP-SRB2* | 0.02 | 0.02 | 0.33 | 0.04 | 0.03 | 0.85 | 0.01 | 0.01 | 0.19 | 0.07 | 0.05 | 1.35 | 0.05 | 0.01 | 0.05 | | 0.15 | 2.78 | 0.06 | 1.09 | 0.05 |
| *SEEP-SRB4* | 0.01 | 0.02 | 1.34 | 0.01 | 0.01 | 1.30 | | 0.01 | 0.39 | | | 0.12 | 0.01 | | 0.01 | | | | 0.03 | 3.28 | 0.01 |
| *Hyd24-12* | 0.04 | 0.03 | 3.44 | 0.01 | 0.02 | 1.15 | | | 0.03 | 0.02 | 0.03 | 1.73 | 0.01 | | 0.01 | | | | | | |
| *Atribacteria* | 0.02 | 0.03 | 1.51 | 0.08 | 0.07 | 4.80 | 0.05 | 0.07 | 3.02 | 0.12 | 0.12 | 7.18 | 0.02 | | 0.02 | | | | | | |
| *Spirochaeta* | | 0.01 | 0.76 | 0.02 | 0.03 | 4.36 | | | | 0.01 | 0.01 | 1.63 | | | 0.01 | | 0.02 | 1.16 | 0.03 | 1.64 | 0.02 |

Table 2: Relative sequence abundances were computed for the top 135 OTUs in the iTag dataset. These OTUs correspond to ~55% of the total sequences in the unfixed bulk sediment. Bacterial and archaeal 16S rRNA gene libraries are included for the core methane seep taxa, with the total number of clones for each library indicated above. Core methane seep taxa were based on Ruff et al. (2015) and include: Candidate Phylum *Atribacteria*, Candidate Division Hyd24-12, *Methanomicrobia*, *Caldilineales*, *Desulfobacterales*, and *Spirochaetales*. While we did recover other *Chloroflexi*, no *Caldilineales* were recovered in iTag or gene library sequencing so they are not included in Table 2. Fixed bulk sediment was chosen for baseline comparison (rather than unfixed) since it includes the potential loss of cells due to fixation and wash steps, thereby processed more similarly to the Magneto-FISH samples. An average and standard deviation for relative sequence abundance among replicates was calculated for each sample set. A ratio of the average relative sequence abundance of Magneto-FISH enrichments compared to the fixed bulk sediment value is reported (Rel. Fixed). Ratios over 1.5 are underlined. 16S rRNA gene bacteria and archaea clone libraries for two Magneto-FISH enrichments and fixed bulk sediment are also included for comparison to recovered iTag diversity.

Table 3: Relative sequence abundances were computed for the top 135 OTUs in the iTag dataset that were not included in the core methane seep microbiome. An average and standard deviation for relative sequence abundance among replicates was calculated for each sample set. A ratio of the average relative sequence abundance of Magneto-FISH enrichments compared to the fixed bulk sediment value is reported (Rel. Fixed). Ratios over 1.5 are underlined. 16S rRNA gene bacteria and archaea clone libraries for two Magneto-FISH enrichments and fixed bulk sediment are also included for comparison to iTag enrichment.

| | Seep1a_1441 | | | DSS_658 | | | Delta_495a | | | MSMX-Eel_932 | | | Fixed Bulk | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Taxon | Avg. | Stdv. | Rel. Fixed | Avg. | Stdv. | Rel. Fixed | Avg. | Stdv. | Rel. Fixed | Avg. | Stdv. | Rel. Fixed | Avg. | Stdv. |
| Desulfarculaceae-uncl | 0.02 | 0.03 | 2.53 | 0.02 | 0.03 | 2.39 | 0.01 | 0.01 | 1.01 | 0.05 | 0.05 | 7.18 | 0.01 | 0.01 |
| Spirochaetae-uncl | | | 0.21 | | | | 0.04 | 0.02 | 3.70 | | | 0.06 | 0.01 | 0.01 |
| Desulfuromusa | 0.05 | 0.05 | 4.17 | | | 0.06 | | | | | 0.01 | 0.39 | 0.01 | |
| Pelobacter | 0.01 | 0.01 | 2.48 | 0.01 | 0.01 | 1.95 | | | 0.10 | | 0.01 | 0.81 | 0.01 | |
| Actinobacteria-OM1 | 0.01 | 0.01 | 0.88 | 0.03 | 0.01 | 2.64 | 0.03 | 0.04 | 2.60 | 0.01 | 0.01 | 0.97 | 0.01 | |
| Alpha-Ancalomicrobium | 0.01 | 0.01 | 2.29 | | | | 0.01 | 0.01 | 2.50 | | | | | |
| Bacteroidetes-Actibacter | 0.01 | 0.02 | 1.38 | | 0.01 | 0.45 | 0.01 | 0.01 | 0.69 | | | 0.06 | 0.01 | |
| Bacteroidetes-BD-2 | 0.03 | 0.01 | 1.49 | 0.01 | 0.02 | 0.58 | 0.02 | 0.01 | 0.94 | 0.03 | 0.02 | 1.29 | 0.02 | |
| Bacteroidetes-Lutibacter | | | | | | | | | | | | | 0.02 | |
| Bacteroidetes-Marinilabiaceae | | | 3.05 | | 0.01 | 3.11 | | | | | | | | |
| Bacteroidetes-SB-1 | | | | | | | | | | | | | 0.01 | |
| Bacteroidetes-SB-5 | 0.01 | 0.01 | 0.89 | 0.01 | 0.01 | 0.96 | | | | | | 0.70 | 0.01 | |
| Bacteroidetes-VC2.1_Bac22 | 0.01 | 0.01 | 0.22 | 0.02 | 0.01 | 0.64 | 0.01 | 0.02 | 0.37 | | | 0.04 | 0.03 | |
| Bacteroidetes-WCHB1-69 | | | 0.29 | | 0.01 | 0.30 | | | | | | 0.11 | 0.01 | 0.01 |
| Chlorobi-PHOS-HE36 | | | | | | | 0.03 | 0.04 | | | | | | |
| Chloroflexi-Anaerolineaceae | 0.02 | 0.02 | 0.73 | 0.01 | 0.01 | 0.43 | 0.01 | 0.01 | 0.23 | 0.02 | 0.02 | 0.69 | 0.03 | 0.01 |
| Chloroflexi-Bellilinea | 0.02 | 0.03 | 4.18 | | | | | | | 0.01 | 0.01 | 2.43 | | |
| Deferribacteres-Caldithrix | 0.01 | 0.01 | 0.31 | 0.01 | 0.01 | 0.19 | | | | 0.01 | 0.01 | 0.46 | 0.03 | |
| Deferribacteres-SAR406 | 0.01 | 0.01 | 3.13 | | | 0.06 | 0.03 | 0.04 | 8.82 | | | 0.18 | | |
| Fibrobacteres-uncl | | 0.01 | 1.50 | | | | 0.01 | | 4.82 | | 0.01 | 1.16 | | |
| Firmicutes-Fusibacter | | | | | | | | | | | | | 0.01 | |
| Firmicutes-Negativicoccus | | | | | | | | | | | | | | |
| Firmicutes-other | | 0.01 | 0.59 | 0.01 | 0.01 | 1.15 | | | | | | | 0.01 | |
| Gam-endosymbionts | | | | | | | | | | 0.01 | 0.01 | 3.28 | | |
| Gamma-other | | | | | | 0.40 | | | | | | 0.34 | | |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| KB1 | | | | | | | | | | | | | | |
| MBGB | | | 0.13 | 0.01 | 0.01 | 1.11 | 0.01 | 0.01 | 1.28 | | | 0.66 | 0.01 | |
| MBGD | | | | 0.01 | 0.01 | 4.48 | 0.01 | 0.01 | 4.89 | | | | | |
| Milano-WF1B-44 | | | | | | | 0.01 | 0.02 | 1.87 | | | 0.02 | 0.01 | |
| OD1 | 0.02 | | 0.88 | 0.03 | 0.03 | 1.20 | 0.01 | 0.01 | 0.43 | 0.03 | 0.01 | 1.16 | 0.02 | |
| Plactomycetes-OM190 | | | | | | | | | | | | | | |
| Planctomycetes-Phycisphaerae | 0.01 | 0.01 | 0.64 | | | | | | | 0.02 | 0.02 | 2.24 | 0.01 | |
| Planctomycetes-Pla4 | | | | | | | | | | | | | | |
| Planctomycetes-SHA-43 | | | | | | | | | | | 0.01 | 1.39 | | |
| Sulfurimonas | | | 0.87 | | | | | | | 0.01 | 0.01 | 1.61 | | |
| Sulfurovum | 0.17 | 0.16 | 1.59 | 0.26 | 0.11 | 2.43 | 0.27 | 0.18 | 2.49 | 0.06 | 0.03 | 0.52 | 0.11 | 0.01 |
| TA06 | | 0.01 | 1.12 | | | | | | | | | | | |
| Thaumarc-uncl | | | | | | 0.12 | | | | | | | 0.01 | |
| Thiohalobacter | | | | 0.01 | 0.01 | | | | | | | | | |
| Thiotrichaceae-uncl | | | | | | | | | | | | | | |
| WS3 | 0.01 | 0.01 | 0.47 | 0.04 | 0.04 | 2.21 | 0.02 | 0.03 | 1.20 | 0.03 | 0.01 | 1.74 | 0.02 | |

than any of the *Deltaproteobacteria* probes compared with the MSMX-Eel_932 probe (Table 4). The most distinct communities were MSMX-Eel_932 enrichment and Delta_495a enrichment, with the highest proportion of unshared branch length (0.97; p-value <0.001). MSMX-Eel_932 enrichment and DSS_658 enrichment had less unshared branch length at 0.88 (<0.001), suggesting MSMX-Eel_932 and DSS_658 probes enrich for a more similar community than MSMX-Eel_932 and Delta_495a probes. Comparison of the MSMX-Eel_932 enrichment and SEEP-1a_1441 enrichment communities was not significant at the <0.001 cutoff. Within the *Deltaproteobacteria* probes, SEEP-1a_1441 enrichment and DSS_658 enrichment had the lowest proportion of unshared community (0.77, <0.001); the most similar community structures were recovered with these two probes. The next lowest proportion of unshared community is between DSS_658 enrichment and Delta_495a enrichment (0.81). SEEP-1a_1441 enrichment and Delta_495a enrichment are least similar, at 0.85. All of these values are highly significant (<0.001). This is consistent with the expectation that the overlap between the target microbial population of the SEEP-1a_1441 probe would be most similar to the target microbial population of the DSS_658 probe, while the Delta_495a enrichment would recover more total *Deltaproteobacteria* diversity.

*Assessing community structure with co-occurrence network analysis*

After determination of statistically significant differences between iTag Magneto-FISH and bulk sediment samples, we computed co-occurrence networks to observe which of the 135 most abundant OTUs were correlated in the methane seep microbial community. By combining the results from 100 separate microbial association calculations, we were able to assign confidence to each microbial association and determine the most robust associations. Significant associations are reported in Sup Table 5 and depicted as a network in Figure 1.

Focusing first on the common ANME syntrophic *Deltaproteobacteria* partner, SEEP-SRB1, this taxon had the most associations in the network including nine positive associations and one negative association (Figure 1). There are two separate sets of SEEP-SRB1 & *Planctomycetes* (AKAU3564 sediment group) positive associations that are both well supported. SEEP-SRB1 is also associated with three other heterotrophic taxa (Candidate Phylum Atribacteria, *Spirochaeta*, and *Bacteroidetes* (VC2.1_Bac22)) and one sulfur-oxidizing taxa (*Sulfurovum*). SEEP-SRB1 was

also associated with Candidate Division Hyd24-12, which has a currently unknown ecophysiology, but could be a heterotroph if the topology of heterotrophic taxa being in the center of the network holds true. Hyd24-12 and *Atribacteria* are also both associated with the second most associated taxa, Candidate Division OD1, but there was no direct association between SEEP-SRB1 and OD1.

SEEP-SRB2 has two of the same associations as SEEP-SRB1 (VC2.1_Bac22 and *Atribacteria*), but is the only *Deltaproteobacteria* associated with MBG-B, *Anaerolineaceae*, and *Desulfoluna* (another *Deltaproteobacteria*). SEEP-SRB4 is associated with *Desulfobulbus*, and the only *Deltaproteobacteria* associated with and ANME (2a/b), WS3, and *Actibacter*. WS3 had high relative sequence abundance in both DSS_658 and MSMX-Eel_932 enrichments (Table 3). *Desulfobulbus* is associated with *Desulfococcus*, the only *Deltaproteobacteria* associated with BD2-2, and SAR406. SAR406 had high relative sequence abundance in Seep1a_1441 and Delta_495a enrichments (Table 3). The heterotroph *Spirochaeta* is also included in the core methane seep microbiome and was associated with *Clostridia* and WS3, in addition to Hyd24-12 and SEEP-SRB1.

In examination of additional OTUs associated with sulfur metabolisms, we found *Sulfurovum* and *Sulfurimonas* (*Epsilonproteobacteria*) were not associated with each other, but are both associated with *Deltaproteobacteria*. *Sulfurimonas* is associated with *Desulfocapsa* and *Sulfurovum* is associated with SEEP-SRB1 and *Desulfobulbus*. *Sulfurovum* had high relative sequence abundance in MSMX-Eel_932 enrichments and *Sulfurimonas* had high relative sequence abundance in Seep-1a_1441, DSS_658, and Delta_495a enrichments (Table 3). The *Gammaproteobacteria, Thiohalobacter,* is only associated with *Anaerolineaceae* and was not elevated in any of the Magneto-FISH enrichments.

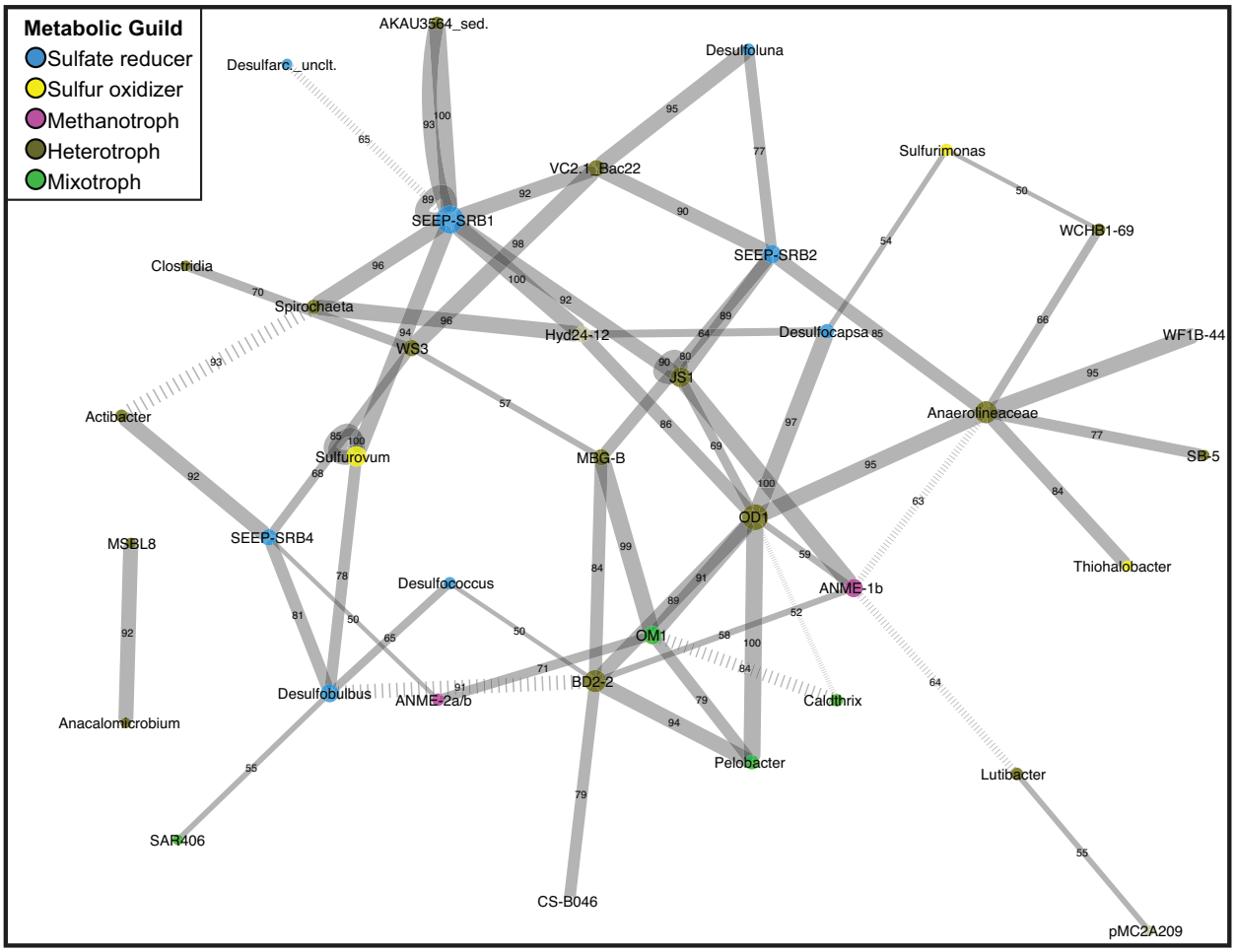
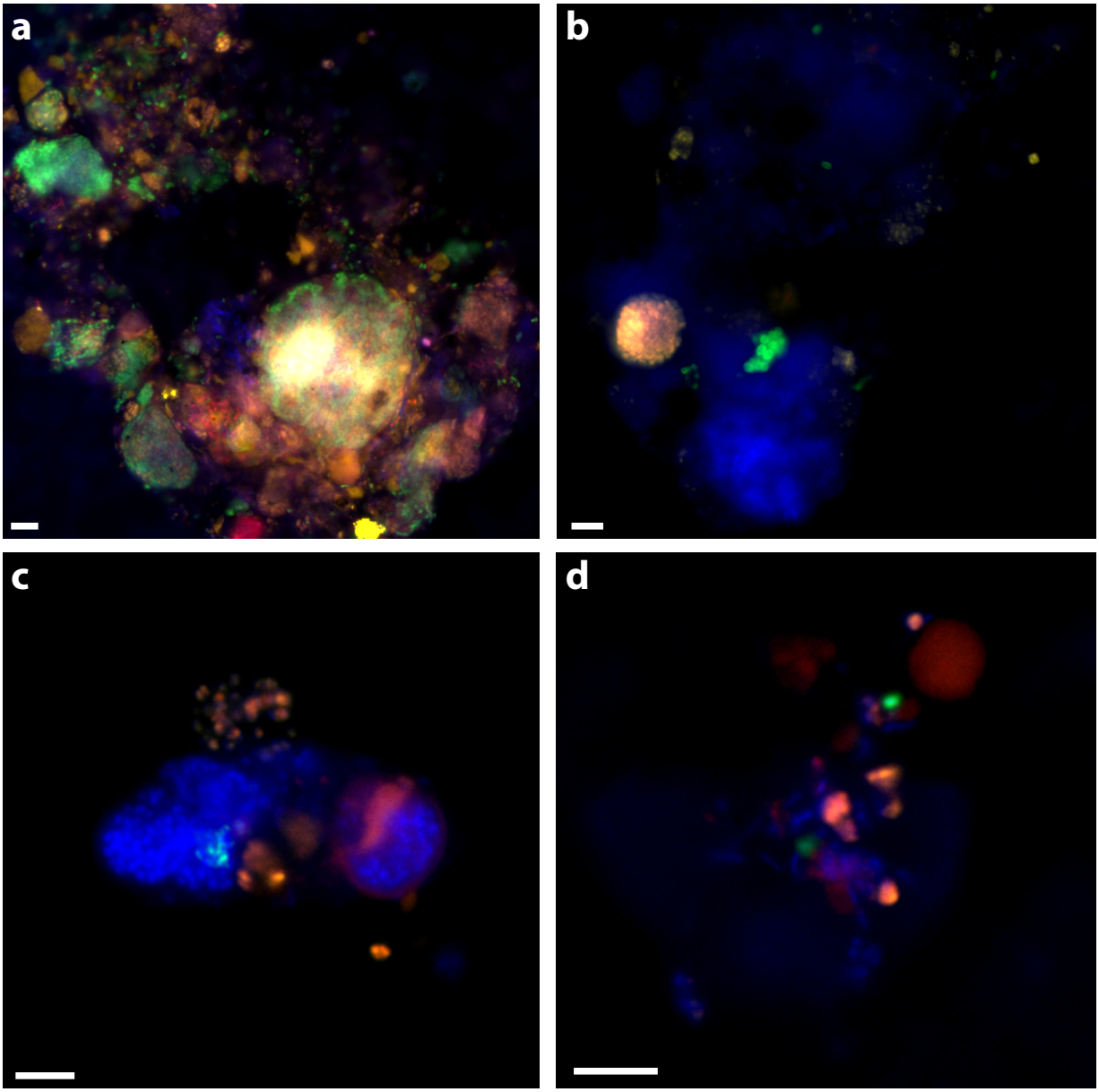



Figure 1: Co-occurrence analysis of the top 135 unique OTUs displayed in network form. Nodes represent the taxonomy of the OTUs in the network and edges are the connections between OTUs. Node size is scaled by number of connecting OTUs and colored by putative metabolic guild (blue – sulfate reducer, yellow – sulfur oxidizer, pink – archaeal methanotroph, brown – heterotroph, green – mixotroph). Edge thickness is scaled by number of occurrences of this association (from 50 to 100 times) and number of occurrences also included along the edge. Negative associations are denoted by hashed lines. The combined network is displayed using Cytoscape, with the average correlation coefficient across all runs determining the distance between nodes and the number of occurrences in 100 network iterations determining edge width. Note JS1 now C.D. *Atribacteria*.

Figure 2: Triple CARD-FISH hybridization using bacterial and archaeal probes targeting DSS_658 (A), Gam42a (B), CF319A/B (C), and Epsi404 (D) in green FITC, with ANME1-350 in red and MSMX-Eel_932 in yellow for all. Scale bar 5 μm for all. DAPI in blue.



| | ANME-1_350 | Eel_932 | DSS_658 | Epsi_404 | Gam_42a | Seep-1a_1441 | CF_319A/B |
|---|---|---|---|---|---|---|---|
| Total | 39 | 70 | 91 | 5 | 12 | 29 | 8 |
| With ANME-1 | | | 36 | 2 | 6 | 21 | 0 |
| With ANME-2 | | | 63 | 1 | 9 | 21 | 4 |
| Percent of all | 39% | 70% | 91% | 10% | 24% | 58% | 16% |
| Percent ANME-1 | | | 36% | 4% | 12% | 42% | 0% |
| Percent ANME-2 | | | 63% | 2% | 18% | 42% | 8% |

Table 4: Aggregate counts from triple CARD-FISH hybridizations with probes targeting ANME-1 (ANME-1_350), all ANME (Eel_932), DSS-type *Deltaproteobacteria* (DSS_658), *Epsilonproteobacteria* (Epsi_404), SEEP-SRB1a (SEEP-1a_1441) and *Cytophaga*, *Bacteroidetes*, *Flavobacterium*, and *Sphingobacterium* (CF_319A/B) associations described in text.

| | Seep1a1441 | DSS658 | Delta495a | Eel932 |
|---|---|---|---|---|
| Seep1a1441 | - | 0.77* | 0.85* | 0.91+ |
| DSS658 | - | - | 0.81* | 0.88* |
| Delta495a | - | - | - | 0.97* |

Table 5: Community comparison of iTag Magneto-FISH samples using weighted UniFrac analysis. Significance of relationship between communities is reported with p-values: *=<0.001, ^=0.002, +=0.030Heterotrophs are the most dominant metabolic guild in the network, and similar to sulfate-reducers, have some of the most connected taxa. The heterotroph OD1 has seven positive correlations, in addition to *Atribacteria* and Hyd24-12 listed above: *Bacteroidetes* (BD2-2), *Actinobacteria* (OM1), *Pelobacter*, ANME-1b, *Chloroflexi* (*Anaerolineaceae)*, and *Desulfocapsa. Anaerolineaceae* and *Bacteroidetes* (BD2-2) both had seven associations, but with different connectivity. BD2-2 was interconnected with other heterotrophs, sulfate-reducers, and archaeal methanotrophs in the main portion of the network, whereas *Anaerolineaceae* was connected to three taxa that share no other connections (two heterotrophs and one *Gammaproteobacteria* sulfur oxidizer). The one other ANME taxa in the network, ANME-1b, is only positively associated with heterotrophs and no known sulfate reducing groups.

*Assessing ANME-bacterial partnerships by CARD-FISH*

To assess ANME and DSS relative cell abundance, 100 aggregates from the same sediment incubation (see Materials & Methods) were analyzed with CARD-FISH and the DSS_658/ANME1-350/MSMX-Eel_932 probe combination. Epsi_404, Gam_42a, SEEP-1a_1441, and CF_319A/B probes were also used with the archaeal probe combination to examine non-DSS bacterial diversity recovered in the network analysis ANME associations. All probes, target populations, and references are listed in Table 1.

30% of aggregates contained an ANME-2 signal (see Materials & Methods; Table 5) and 39% of aggregates had an ANME-1 signal. ANME-1 and ANME-2 identified cells were also consistent with expected morphologies. Multiple clusters of mixed-type ANME/DSS, DSS-only, ANME-only, DSS/non-ANME, and non-DSS/non-ANME aggregates were observed with the ANME-1_350, MSMX-Eel_932, and DSS_658 probe combination (Figure 2a). There were no clear examples of aggregates with ANME/non-DSS hybridized cells, though we found many instances where both ANME and non-DSS cells were part of a larger aggregate cluster with other cell types. ANME-1 cells often occurred in the matrix surrounding tightly clustered ANME-2 aggregates. The SEEP-1a_1441 probe, targeting a subgroup of DSS, was observed to hybridize with aggregates that contained ANME-1 and ANME-2 cells, but usually with SEEP-SRB1/ANME-2 in tight association and ANME-1 in more peripheral association. Five of the SEEP-SRB1/ANME-2 aggregates did not have ANME-1 cells (10%) and three of the SEEP-SRB1/ANME-1 aggregates did not have ANME-2 cells.

Ten percent of aggregates (n=50 counted) hybridized with the Epsi_404 probe, broadly targeting members of the *Epsilonproteobacteria*. These *Epsilonproteobacteria* were mostly found in association with other bacteria and occasionally, loosely associated with some ANME. Epsi_404 hybridized cells were generally ovoid and scattered throughout an EPS matrix of cells, as depicted in Figure 2d. There was no apparent preference for *Epsilonproteobacteria* association with ANME-1 or ANME-2 aggregates (Table 4). A higher percentage of aggregates had *Gammaproteobacteria* cells (24% of 50) than *Epsilonproteobacteria* cells, and there was a slightly higher co-occurrence with ANME-2 (18%) than ANME-1 (12%) hybridized. The dominant *Gammaproteobacteria* morphology observed was a cluster or chain of large (~1 μm) ovoid cells. Gam_42a hybridizing cell clusters and chains were found both separately and associated with other bacteria, as in Figure 2b, where they are predominately an unidentified cluster stained by DAPI with a sub-aggregate of ANME-2 cells. CF319A and CF319B were used to target *Cytophaga*, *Bacteroidetes*, *Flavobacterium*, and *Sphingobacterium*. Eight percent (n=50 counted) of aggregates contained cells positively hybridizing with the CFB probe, generally observed as clustered filaments or rods (Figure 2c). Half of these aggregates also had ANME-2 hybridized cells. No CFB cells were observed to co-associate with ANME-1.

**Discussion.**

*Evaluation of Magneto-FISH with iTag*

Challenges accompanying downstream analysis of Magneto-FISH enrichments are primarily associated with low DNA yield and poor DNA quality from aldehyde fixation (for further discussion of fixation effects see Trembath-Reichert et al. 2013). Low template concentration exacerbates amplification of contaminating sequences since target and non-target templates can approach parity in a PCR reaction. Low template concentration has also been shown to create random variation in amplification products in dilution experiments (Chandler et al. 1997), which could explain the high variation seen in Magneto-FISH enrichment relative sequence abundances compared to bulk sediment samples. Despite these challenges, the DNA recovered from Magneto-FISH enrichments has been shown to increase the sequence abundance of target organisms relative to the bulk sediment by 16S rRNA gene sequencing and metagenomics on various Next Generation sequencing platforms (Pernthaler et al. 2008; Trembath-Reichert et al. 2013). In this study, conventional cloning and sequencing of full-length bacterial and archaeal 16S rRNA genes had fewer contamination issues as compared to iTag sequencing with universal primers. Our Magneto-FISH experiments were designed to mitigate as many sampling and iTag sequencing biases between samples as possible, by concurrently extracting, amplifying, and sequencing all Magneto-FISH samples in parallel, including biological and technical replicates. The relative ratio of contaminant reads to environmental OTU's were higher in Magneto-FISH enrichments than in bulk sediment samples, but bulk sediment could be used to separate indigenous community members from putative contaminants in the Magneto-FISH samples (see Materials & Methods). This provided a conservative Magneto-FISH dataset for statistical analyses and demonstrated the importance of parallel processing sequencing of bulk and separated samples.

In addition to issues with contaminating sequences, we also observed bias against some core methane seep microbiome taxa, where these taxa were consistently underrepresented by iTag when compared to gene libraries and CARD-FISH. ANME-2 was the most underrepresented taxon in iTag sequencing of the bulk sediment and mock communities, with much greater relative sequence and relative cell abundance in gene library sequencing and CARD-FISH analysis, respectively. It is most likely that iTag sequencing bias with the EMP primer set is the reason ANME-2c was not enriched in the Magneto-FISH samples and absent from microbial community network analysis. Members of the ANME-2a/b were also, to a lesser extent, underrepresented with iTag. In addition

to our gene libraries and CARD-FISH analysis, independent assays using FISH with mono labeled oligonucleotide probes from this sediment incubation further confirmed the abundance of ANME-2 aggregates; 25% of aggregates were ANME-2c and 17% of aggregates were ANME-2b, with about half of ANME-2 aggregates associating with a bacterial partner other than SEEP-SRB1 (Supplement McGlynn et al. 2015). We conclude that while expected ANME-2 associations were not recovered, they can be explained by EMP iTag bias and therefore do not reduce the validity of other non-ANME-2 associations recovered in the co-occurrence analysis (see Sup Table 2 and 3 captions for further discussion of ANME-2c bias). Although ANME-1a was not underrepresented in the iTag data, it still does not appear in the co-occurrence network. In other co-occurrence network studies dominant OTUs were not associated with the majority of the microbial community, which was thought to be due to a high degree of functional redundancy (Mu & Moreau 2015). Possible functional redundancy with other archaeal groups, or simply non-specific, loose spatial association with many taxa, as suggested by CARD-FISH analysis, could explain why ANME-1a was not recovered in our network analysis.

Despite this unanticipated methodological bias, iTag sequencing is a valid and valuable tool when combined with Magneto-FISH enrichment techniques for microbial association hypothesis development and testing. For example, we saw more bacterial OTUs, especially among *Deltaproteobacteria*, in the iTag samples compared with conventional gene libraries and the core methane seep taxon Hyd24-12 was not even observed among gene library sequences.

*Magneto-FISH enrichment*

This study provides a novel combination of nested Magneto-FISH enrichments and microbial community network analysis methods to develop hypotheses regarding specific lineage associations and, by inference, discusses the potential for additional metabolic interactions relating to sulfur cycling in methane seep sediments. Notwithstanding the low recovery of ANME-2 OTUs, there was statistical support for Magneto-FISH enrichments increasing the relative iTag sequence abundance of target organisms. Statistical analyses demonstrated SEEP-SRB1 and *Desulfobulbus* OTUs were significantly different in Magneto-FISH samples (t-tests), and these OTUs were significantly more enriched in Magneto-FISH samples using linear discriminant analysis (LDA) effect size (LEfSe). Additionally, weighted UniFrac analysis showed the highest

percentage of shared phylogeny was between the clade-specific SEEP-1a_1441 probe and the family-specific *Desulfobacteraceae* DSS_658 probe enrichments. Therefore these Magneto-FISH samples contain microbial community overlap consistent with probe target specificity, even when some dominant community members are not represented at expected relative sequence abundance in the iTag analysis (ANME-2).

Magneto-FISH enrichment relative sequence abundance followed expected trends for *Deltaproteobacteria* (Table 2). SEEP-SRB1 had the highest relative sequence abundance in Seep-1a_1441 and MSMX-Eel_932 enrichments, which should target this group. *Desulfobulbus* had the highest relative sequence abundance in the Delta_495a enrichment, which was the only Magneto-FISH probe that should hybridize to this group (though *Desulfobulbus* could also be retrieved via association with other target organisms). OTUs affiliated with *Desulfoluna* (within the *Desulfobacteraceae*) had the highest relative sequence abundance of all *Deltaproteobacteria* in the DSS_658 enrichment and are also targeted by the DSS_658 probe. *Desulfoluna* were not specifically targeted by MSMX-Eel_932 or Seep-1a_1441 probes, but had high relative sequence abundane in these samples and may have a potential association with ANME/DSS consortia. Also, *Atribacteria* (JS1) was recovered in all iTag sequencing of Magneto-FISH enrichments, suggesting they may associate with either DSS/ANME or DSB/ANME consortia. Members of the Hyd24-12 were only recovered in Seep1a_1441 and MSMX-Eel_932 enrichments and may preferentially associate with SEEP-SRB1a/ANME consortia.

Evaluating our iTag relative sequence abundance data with co-occurrence analysis, we developed hypotheses that were not subject to the variation between Magneto-FISH enrichment replicates; associated taxa should always co-vary, even when they are less abundant than expected. Within the core methane seep taxa, high relative sequence abundances of *Atribacteria* and Hyd24-12 with SEEP-SRB1 targeting Magneto-FISH enrichments were upheld by the network. Hyd24-12 is highly associated with SEEP-SRB1, whereas *Atribacteria* is highly associated with both SEEP-SRB1 (DSS) and SEEP-SRB2 (DSB). While *Atribacteria* have not been cultured, metagenomic sequencing suggests they are likely heterotrophic anaerobes involved in fermentation (Nobu et al. 2015). Hyd24-12 was first cloned from Hydrate Ridge (Knittel et al. 2003) and has been cited as a core methane seep microbial taxon (Ruff et al. 2015), but nothing is known about its physiology. The Hyd24-12/SEEP-SRB1 association was also one of the four unique associations that were

recovered in all the network computations (n=100). These results may aid in determining a role for these enigmatic candidate phyla of the methane seep microbiome.

*Methanomicrobia* and *Deltaproteobacteria* only had one co-occurrence in our network. The one statistically supported network ANME/SRB association was between ANME-2a/b and SEEP-SRB4. SEEP-SRB4, belonging to the *Desulfobulbaceae* (Knittel et al. 2003), and ANME-2a/b both had high relative sequence abundance in the ANME-targeting MSMX-Eel_932 enrichment bacterial 16S rRNA gene library. There have been FISH-confirmed physical associations between ANME-2/ANME-3 and *Desulfobulbaceae* (Green-Saxena et al. 2014; Löesekann et al. 2007; Pernthaler et al. 2008) in AOM systems. SEEP-SRB4 was also strongly associated with the candidate phyla WS3 in the network, and WS3 was enriched in both DSS_658 and MSMX-Eel_932 enrichments. Both SEEP-SRB4 associations with ANME-2a/b and WS3 warrant future study.

While expected ANME-2/*Deltaproteobacteria* associations were not recovered (see *Evaluation of Magneto-FISH with iTag*), network analysis did recover many *Deltaprotobacteria* co-occurring with bacterial groups. Almost half of all positive associations contained a *Deltaproteobacteria* OTU (30/61), suggesting a dominant role for the sulfur cycle metabolisms. Of those, 21 associations were with a non-*Proteobacteria* OTU including a number of candidate organisms as described above. The association between SEEP-SRB1 and 'AKAU3564,' a *Planctomycetes*-affiliated heterotrophic sediment group, was observed twice with two separate OTU associations in this clade that were both strongly supported (occurring 100/100 and 93/100 times, respectively, that the network analysis was run, Sup Table 5). This Planctomycete group was first described in methane hydrate bearing deep marine sediments of the Peru Margin (Inagaki et al. 2006). *Planctomycetes*-associated sequences were previously recovered in association with ANME-2c Magneto-FISH samples from the Eel River Basin, where the preferred partner was observed to be the SEEP-SRB1 group (Pernthaler et al. 2008). It follows that SEEP-SRB1 may also co-occur with *Planctomycetes*, if these organisms are affiliated (either directly or indirectly) with ANME-2 consortia. By similar logic, although it did not have high relative sequence abundance in the Seep1a_1441 enrichment, this could explain the high relative sequence abundance of this group in the MSMX-Eel_932 enrichment (Table 3). *Planctomycetes* targeted CARD-FISH hybridization using the general *Planctomycetes* probe Pla_886 was attempted; however, many cells with a

morphology similar to ANME-1 were hybridized and the results were deemed inconclusive. This ambiguity could be due to the probe's single base pair mismatch to 97% of ANME-1a, 94% of ANME-1b, and 25% of ANME-2b, even if this mismatch was centrally located (SILVA TestProbe online tool, Greuter et al. 2015). *Spirochaeta* was also associated with SEEP-SRB1, in addition to Hyd24-12 and WS3, and had high relative sequence abundance in both the DSS_658 and MSMX-Eel_932 enrichments (Table 2). In addition to being core methane seep microbial taxa, some members of the *Spirochaetes* have sulfide-oxidizing capabilities in mats with sulfidogenic bacteria (Dubinina et al. 2004) and it is possible that these organisms may be utilizing sulfide produced in seep systems as well.

*Epsilonproteobacteria* and *Deltaproteobacteria* were the most common intra-*Proteobacteria* association in the network and have been shown to co-occur in many sulfidic habitats (Campbell et al. 2006; Omoregie et al. 2008), where *Epsilonproteobacteria* oxidize sulfur and *Deltaproteobacteria* disproportionate or reduce sulfur species (Pjevac et al. 2014). In the network, *Sulfurovum* was associated with both SEEP-SRB1 and *Desulfobulbus*, and this was also seen in the relative sequence abundance data where *Sulfurovum* had high relative sequence abundance in all of the *Deltaproteobacteria* Magneto-FISH enrichments. *Epsilonproteobacteria* have been shown to oxidize sulfide to $S^{o}$ or $HS^{-}$ to sulfate in microbial mats (Pjevac et al. 2014), allowing some sulfur substrate differentiation between these *Epsilonproteobacteria* groups in this system. *Sulfurimonas* was not strongly associated with any *Deltaproteobacteria* in the network analysis and only had high relative sequence abundance in the MSMX-Eel_932 enrichment (16S rRNA gene iTag, 16S rRNA gene bacterial, and soxB gene libraries; see Sup Figure 2 for further discussion of metabolic genes). CARD-FISH analysis using probe Epsi_404 confirmed the presence of *Epsilonproteobacteria* cells within some ANME and other non-hybridized cell-containing loose aggregates, but did not appear to be in the tight physical association characteristic of ANME/SRB consortia. While cultured representatives of these *Epsilonproteobacteria* have optimum growth with some oxygen present (Inagaki et al. 2003; Inagaki et al. 2004), it is possible that these uncultured methane seep *Epsilonproteobacteria* may be able to use other oxidants such as nitrate or intermediate sulfur species while in anaerobic incubation conditions.

In comparison to *Delta*- and *Epsilonproteobacteria*, there was only one *Gammaproteobacteria* OTU in the network (*Thiohalobacter*, with one *Anaerolineaceae* association). Cultured

representatives of *Thiohalobacter* have diverse sulfur capabilities, including thiocyanate metabolism, but are not known to form associations with other sulfur cycling organisms (Sorokin et al. 2010). This differentiation between *Gamma*- and *Epsilon*-/*Deltaproteobacteria* has been seen in other systems such as sulfidic cave biofilms (Macalady et al. 2008) or in microbial mats on marine sediments (Pjevac et al. 2014). Gam_42a hybridizing cells (*Gammaproteobacteria*) were observed to form aggregates with non-ANME and non-*Desulfobulbaceae* (DSS) cells in our CARD-FISH analysis, but the identity of these organisms was not determined. While not recovered in the network, the majority of the *Gammaproteobacteria* OTUs observed by iTag from the both the bulk sediment and MSMX-Eel_932 Magneto-FISH 16S rRNA gene (Table 1) and aprA gene libraries (see Sup Figure 2 for further discussion of metabolic genes) were from the SILVA taxonomy endosymbiont clade. This endosymbiont clade houses organisms with a carbon-fixation/sulfur-oxidation metabolism (Duperron et al. 2012; Goffredi 2010) and is predicted to be an important member of the sulfur and carbon cycles in marine sediments outside of an endosymbiotic lifestyle (Lenk et al. 2011).

There were also three unique, positive *Deltaproteobacteria-Deltaproteobacteria* associations observed in the network: *Desulfobulbus*/*Desulfococcus*, *Desulfobulbus*/SEEP-SRB4, *Desulfoluna*/SEEP-SRB2. These multiple intra-*Deltaproteobacteria* associations suggest there may be further nuances to be explored in the *Deltaproteobacteria* community structure, perhaps akin to the nitrate based partitioning observed between DSB and DSS in seep sediments (Green-Saxena et al. 2014). *Desulfobulbus* was also associated with SAR406, and SAR406 had high relative sequence abundance in the Delta495a enrichments. SAR406 (Marine Group A) fosmids contained polysulfide reductase genes that may be used for dissimilatory polysulfide reduction (Wright et al. 2014). *Desulfobulbus* can also use polysulfide, in addition to a range of other sulfur sources (Fuseler & Cypionka 1995), potentially linking these two taxa.

**Conclusions.**

Our findings support the utilization of paired Magneto-FISH and iTag sequencing in developing and testing hypotheses to interrogate complex interactions in microbial communities. Contaminants and amplification bias can be identified and mitigated with diversity assessment by multiple means (i.e. multiple iTag primer sets, FISH surveys, or non-16S rRNA gene surveys) and

parallel processing of control samples (bulk sediment and no-template) along with Magneto-FISH enrichments. Since it may not always be known *a priori* which taxa are in an environmental sample, sequencing of a defined mock community may not be an option for assessing bias. However, in our case, prior knowledge of major seep taxa enabled assessment of amplification bias by iTag. It should also be noted that the degree of bias was more pronounced in the environmental samples than our mock samples; therefore mock community samples may not fully capture the degree of bias, but can be useful in identifying which taxa may be the most biased. We found the bulk sediment 16S rRNA gene libraries to be the most useful for determining which of the most abundant taxa were affected by amplification bias. Future studies may benefit more from bulk sediment analysis by a range of iTag primer sets or gene libraries to assess potential sequencing biases in a new microbial community.

Multiple statistical methods supported differences between Magneto-FISH enrichments and the bulk sediment. We also found variation between SparCC network computations. Therefore, we added confidence to network associations by reporting the number of times an association was recovered out of 100 co-occurrence iterations along with correlation and p-value.

Our resultant microbial community network had many statistically significant methane seep taxa correlations beyond the common ANME/SRB association. The downplay of anaerobic methanotrophs in our iTag sequencing may have had the beneficial effect of bringing fermenters to the forefront, highlighting their complex role in methane seep microbial communities. Within the core methane seep microbiome taxa, there were strong associations between *Atribacteria* and Hyd24-12 and *Deltaproteobacteria*, but no direct association between *Atribacteria* and Hyd24-12. This may indicate a different niche for these two currently uncultured groups in methane seep systems. *Sulfurovum* and *Sulfurimonas* were differentiated as either *Deltaproteobacteria*-associated or archaea-associated, respectively. There were statistically significant associations between *Deltaproteobacteria* and non-*Proteobacteria*, such as the *Planctomycetes* sediment group 'AKAU3564,' and groups that contained neither SRB nor ANME but had high statistical significance, such as MBG-B and OM1. Future development and application of more specific FISH probes will assist in further hypotheses development and testing of these associations in Hydrate Ridge methane seeps.

Some groups, such as *Gammaproteobacteria*, appeared to have associations with other microbes based on broad FISH surveys and Magneto-FISH relative sequence abundance data, but were not recovered in the network analysis. Determination of the specific *Gammaproteobacteria* involved in associations via FISH probe development or other means (Hatzenpichler et al., in review) will also aid in refining why associations might be missed in the microbial network analysis based on DNA taxa co-occurrence. In summary, a continual feedback loop between microbial identification and isolation techniques and gene based statistical analyses is required to tease apart interactions within complex microbial systems. The combination of Magneto-FISH and high throughput, parallel iTag sequencing provides an effective bridge between these two modes.

**Acknowledgements**.

We thank the crew of the R/V *Atlantis* and DSV *JASON II*, as well as Abigail Green-Saxena and Joshua Steele for assistance with optimization of the Magneto-FISH protocol and Stephanie Connon for assistance with sequencing. We also are grateful to Katherine Dawson, Emil Ruff, and two anonymous reviewers for providing comments on this manuscript.

**References**:


Akerman NH, Butterfield DA, and Huber JA. 2013. Phylogenetic diversity and functional gene patterns of sulfur-oxidizing subseafloor Epsilonproteobacteria in diffuse hydrothermal vent fluids. *Frontiers in Microbiology* 4:185. 10.3389/fmicb.2013.00185

Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, and Stahl DA. 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Applied and Environmental Microbiology* 56:1919-1925.

Ashkin A. 1997. Optical trapping and manipulation of neutral particles using lasers. *Proceedings of the National Academy of Sciences* 94:4853-4860.

Barberán A, Bates ST, Casamayor EO, and Fierer N. 2012. Using network analysis to explore co-occurrence patterns in soil microbial communities. *The ISME Journal* 6:343-351. 10.1038/ismej.2011.119

Berry D, Mader E, Lee TK, Woebken D, Wang Y, Zhu D, Palatinszky M, Schintlmeister A, Schmid MC, Hanson BT, Shterzer N, Mizrahi I, Rauch I, Decker T, Bocklitz T, Popp Jr, Gibson CM, Fowler PW, Huang WE, and Wagner M. 2015. Tracking heavy water (D2O) incorporation for identifying and sorting active microbial cells. *Proceedings of the National Academy of Sciences* 112:E194-E203. 10.1073/pnas.1420406112

Blazejak A, Kuever J, Erséus C, Amann R, and Dubilier N. 2006. Phylogeny of 16S rRNA, Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase, and Adenosine 5'-Phosphosulfate Reductase Genes from Gamma- and Alphaproteobacterial Symbionts in Gutless Marine Worms (Oligochaeta) from Bermuda and the Bahamas. *Applied and Environmental Microbiology* 72:5527-5536. 10.1128/aem.02441-05

Boetius A, Ravenschlag K, Schubert CJ, Rickert D, Widdel F, Gieseke A, Amann R, Jorgensen BB, Witte U, and Pfannkuche O. 2000. A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature* 407:623-626.

Campbell BJ, Engel AS, Porter ML, and Takai K. 2006. The versatile [epsi]-proteobacteria: key players in sulphidic habitats. *Nat Rev Micro* 4:458-468.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, and Knight R. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6:1621-1624.

Chaffron S, Rehrauer H, Pernthaler J, and von Mering C. 2010. A global network of coexisting microbes from environmental and whole-genome sequence data. *Genome Research* 20:947-959. 10.1101/gr.104521.109

Chandler DP, Fredrickson JK, and Brockman FJ. 1997. Effect of PCR template concentration on the composition and distribution of total community 16S rDNA clone libraries. *Molecular Ecology* 6:475-482.

DeLong EF. 1992. Archaea in coastal marine environments. *Proc Natl Acad Sci U S A* 89:5685-5689.

Dubinina GA, Grabovich MY, and Chernyshova YY. 2004. The role of oxygen in the regulation of the metabolism of aerotolerant spirochetes, a major component of ‚ÄúThiodendron‚Äù bacterial sulfur mats. *Microbiology* 73:621-628. 10.1007/s11021-005-0001-3

Duperron S, Rodrigues CF, Léger N, Szafranski K, Decker C, Olu K, and Gaudron SM. 2012. Diversity of symbioses between chemosynthetic bacteria and metazoans at the Guiness cold seep site (Gulf of Guinea, West Africa). *MicrobiologyOpen* 1:467-480. 10.1002/mbo3.47

Edgar RC, Haas BJ, Clemente JC, Quince C, and Knight R. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27:2194-2200. 10.1093/bioinformatics/btr381

Friedman J, and Alm EJ. 2012. Inferring correlation networks from genomic survey data.

Fuseler K, and Cypionka H. 1995. Elemental sulfur as an intermediate of sulfide oxidation with oxygen byDesulfobulbus propionicus. *Archives of Microbiology* 164:104-109. 10.1007/bf02525315

Goffredi SK. 2010. Indigenous ectosymbiotic bacteria associated with diverse hydrothermal vent invertebrates. *Environmental Microbiology Reports* 2:479-488. 10.1111/j.1758-2229.2010.00136.x

Green-Saxena A, Dekas AE, Dalleska NF, and Orphan VJ. 2014. Nitrate-based niche differentiation by distinct sulfate-reducing bacteria involved in the anaerobic oxidation of methane. *ISME J* 8:150-163. 10.1038/ismej.2013.147

Greuter D, Loy A, Horn M, and Rattei T. 2015. probeBase,Äîan online resource for rRNA-targeted oligonucleotide probes and primers: new features 2016. *Nucleic Acids Research*. 10.1093/nar/gkv1232

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, and Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59:307-321. 10.1093/sysbio/syq010

Hatzenpichler R, and Orphan V. 2015. Detection of Protein-Synthesizing Microorganisms in the Environment via Bioorthogonal Noncanonical Amino Acid Tagging (BONCAT). Totowa, NJ: Humana Press, 1-13.

Hatzenpichler R, Scheller S, Tavormina PL, Babin BM, Tirrell DA, and Orphan VJ. 2014. In situ visualization of newly synthesized proteins in environmental microbes using amino acid tagging and click chemistry. *Environmental Microbiology* 16:2568-2590. 10.1111/1462-2920.12436

House CH, Orphan VJ, Turk KA, Thomas B, Pernthaler A, Vrentas JM, and Joye SB. 2009. Extensive carbon isotopic heterogeneity among methane seep microbiota. *Environmental Microbiology* 11:2207-2215. 10.1111/j.1462-2920.2009.01934.x

Inagaki F, Nunoura T, Nakagawa S, Teske A, Lever M, Lauer A, Suzuki M, Takai K, Delwiche M, Colwell FS, Nealson KH, Horikoshi K, D,ÄôHondt S, and J√∏rgensen BB. 2006. Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin. *Proceedings of the National Academy of Sciences of the United States of America* 103:2815-2820. 10.1073/pnas.0511033103

Inagaki F, Takai K, Kobayashi H, Nealson KH, and Horikoshi K. 2003. Sulfurimonas autotrophica gen. nov., sp. nov., a novel sulfur-oxidizing Œμ-proteobacterium isolated from hydrothermal sediments in the Mid-Okinawa Trough. *International Journal of Systematic and Evolutionary Microbiology* 53:1801-1805. 10.1099/ijs.0.02682-0

Inagaki F, Takai K, Nealson KH, and Horikoshi K. 2004. Sulfurovum lithotrophicum gen. nov., sp. nov., a novel sulfur-oxidizing chemolithoautotroph within the Œμ-Proteobacteria isolated from Okinawa Trough hydrothermal sediments. *International Journal of Systematic and Evolutionary Microbiology* 54:1477-1482. 10.1099/ijs.0.03042-0

Kalyuzhnaya M, Lidstrom M, and Chistoserdova L. 2008. Real-time detection of actively metabolizing microbes by redox sensing as applied to methylotroph populations in Lake Washington. *The ISME Journal* 2:696-706.

Knittel K, Boetius A, Lemke A, Eilers H, Lochte K, Pfannkuche O, Linke P, and Amann R. 2003. Activity, distribution, and diversity of sulfate reducers and other bacteria in sediments above gas hydrate (Cascadia Margin, Oregon). *Geomicrobiology Journal* 20:269-294.

Kozich JJ, Westcott SL, Baxter NT, Highlander SK, and Schloss PD. 2013. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Applied and Environmental Microbiology* 79:5112-5120. 10.1128/aem.01043-13

Lane DJ. 1991. 16S/23S rRNA sequencing. In: Stackebrandt EaG, M., ed. *Nucleic acid techniques in bacterial systematics*. Chichester, England: John Wiley & Sons, 115-175.

Lenk S, Arnds J, Zerjatke K, Musat N, Amann R, and Mussmann M. 2011. Novel groups of Gammaproteobacteria catalyse sulfur oxidation and carbon fixation in a coastal, intertidal sediment. *Environ Microbiol* 13:758-774. 10.1111/j.1462-2920.2010.02380.x

Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, Park YM, Buso N, and Lopez R. 2015. The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Research*. 10.1093/nar/gkv279

Löesekann T, Knittel K, Nadalig T, Fuchs B, Niemann H, Boetius A, and Amann R. 2007. Diversity and abundance of aerobic and anaerobic methane oxidizers at the Haakon Mosby Mud Volcano, Barents Sea. *Applied and Environmental Microbiology* 73:3348-3362.

Loy A, Lehner A, Lee N, Adamczyk J, Meier H, Ernst J, Schleifer K-H, and Wagner M. 2002. Oligonucleotide Microarray for 16S rRNA Gene-Based Detection of All Recognized Lineages of Sulfate-Reducing Prokaryotes in the Environment. *Applied and Environmental Microbiology* 68:5064-5081. 10.1128/aem.68.10.5064-5081.2002

Lozupone C, and Knight R. 2005. UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. *Applied and Environmental Microbiology* 71:8228-8235. 10.1128/aem.71.12.8228-8235.2005

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, Buchner A, Lai T, Steppi S, Jobb G, Förster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Jost R, König A, Liss T, Lüßmann R, May M, Nonhoff B, Reichel B, Strehlow R, Stamatakis A, Stuckmann N, Vilbig A, Lenke M, Ludwig T, Bode A, and Schleifer K-H. 2004. ARB: a software environment for sequence data. *Nucleic Acids Research* 32:1363-1371. 10.1093/nar/gkh293

Macalady JL, Dattagupta S, Schaperdoth I, Jones DS, Druschel GK, and Eastman D. 2008. Niche differentiation among sulfur-oxidizing bacterial populations in cave waters. *ISME J* 2:590-601.

Macalady JL, Lyon EH, Koffman B, Albertson LK, Meyer K, Galdenzi S, and Mariani S. 2006. Dominant Microbial Populations in Limestone-Corroding Stream Biofilms, Frasassi Cave System, Italy. *Applied and Environmental Microbiology* 72:5596-5609. 10.1128/aem.00715-06

Malfatti F, and Azam F. 2010. Atomic force microscopy reveals microscale networks and possible symbioses among pelagic marine bacteria. *Aquatic Microbial Ecology* 58:1-14. doi: 10.3354/ame01355

Manz W, Amann R, Ludwig W, Vancanneyt M, and Schleifer K-H. 1996. Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum cytophaga-flavobacter-bacteroides in the natural environment. *Microbiology* 142:1097-1106. doi:10.1099/13500872-142-5-1097

Manz W, Amann R, Ludwig W, Wagner M, and Schleifer KH. 1992. PHYLOGENETIC OLIGODEOXYNUCLEOTIDE PROBES FOR THE MAJOR SUBCLASSES OF PROTEOBACTERIA - PROBLEMS AND SOLUTIONS. *Systematic and Applied Microbiology* 15:593-600.

Manz W, Eisenbrecher M, Neu TR, and Szewzyk U. 1998. Abundance and spatial organization of Gram-negative sulfate-reducing bacteria in activated sludge investigated by in situ probing with specific 16S rRNA targeted oligonucleotides. *Fems Microbiology Ecology* 25:43-61. 10.1111/j.1574-6941.1998.tb00459.x

McGlynn SE, Chadwick GL, Kempes CP, and Orphan VJ. 2015. Single cell activity reveals direct electron transfer in methanotrophic consortia. *Nature* 526:531-535. 10.1038/nature15512

Melin J, and Quake SR. 2007. Microfluidic large-scale integration: the evolution of design rules for biological automation. *Annu Rev Biophys Biomol Struct* 36:213-231.

Mu A, and Moreau JW. 2015. The Geomicrobiology of CO2 geosequestration: a focused review on prokaryotic community responses to field-scale CO2 injection. *Frontiers in Microbiology* 6. 10.3389/fmicb.2015.00263

Neef A, Amann R, Schlesner H, and Schleifer K-H. 1998. Monitoring a widespread bacterial group: in situ detection of planctomycetes with 16S rRNA-targeted probes. *Microbiology* 144:3257-3266. doi:10.1099/00221287-144-12-3257

Niemann H, Loesekann T, de Beer D, Elvert M, Nadalig T, Knittel K, Amann R, Sauter EJ, Schluter M, and Klages M. 2006. Novel microbial communities of the Haakon Mosby mud volcano and their role as a methane sink. *Nature* 443:854-858.

Nobu MK, Dodsworth JA, Murugapiran SK, Rinke C, Gies EA, Webster G, Schwientek P, Kille P, Parkes RJ, Sass H, Jorgensen BB, Weightman AJ, Liu W-T, Hallam SJ, Tsiamis G, Woyke T, and Hedlund BP. 2015. Phylogeny and physiology of candidate phylum /`Atribacteria/' (OP9/JS1) inferred from cultivation-independent genomics. *ISME J*. 10.1038/ismej.2015.97

Omoregie EO, Mastalerz V, de Lange G, Straub KL, Kappler A, R√∏y H, Stadnitskaia A, Foucher J-P, and Boetius A. 2008. Biogeochemistry and Community Composition of Iron- and Sulfur-Precipitating Microbial Mats at the Chefren Mud Volcano (Nile Deep Sea Fan, Eastern Mediterranean). *Applied and Environmental Microbiology* 74:3198-3215. 10.1128/aem.01751-07

Orphan VJ. 2009. Methods for unveiling cryptic microbial partnerships in nature. *Current Opinion in Microbiology* 12:231-237.

Orphan VJ, Hinrichs KU, Ussler W, Paull CK, Taylor LT, Sylva SP, Hayes JM, and Delong EF. 2001a. Comparative analysis of methane-oxidizing archaea and sulfate-reducing bacteria in anoxic marine sediments. *Applied and Environmental Microbiology* 67:1922-1934.

Orphan VJ, House CH, Hinrichs K-U, McKeegan KD, and DeLong EF. 2001b. Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis. *Science* 293:484-487.

Orphan VJ, House CH, Hinrichs KU, McKeegan KD, and DeLong EF. 2002. Direct phylogenetic and isotopic evidence for multiple groups of archaea involved in the anaerobic oxidation of methane. *Geochimica Et Cosmochimica Acta* 66:A571-A571.

Parada A, Needham DM, and Fuhrman JA. 2015. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time-series and global field samples. *Environmental Microbiology*:n/a-n/a. 10.1111/1462-2920.13023

Pernthaler A, Dekas AE, Brown CT, Goffredi SK, Embaye T, and Orphan VJ. 2008. Diverse syntrophic partnerships from deep-sea methane vents revealed by direct cell capture and metagenomics. *Proceedings of the National Academy of Sciences* 105:7052-7057. 10.1073/pnas.0711303105

Pjevac P, Kamyshny A, Dyksma S, and Mußmann M. 2014. Microbial consumption of zero-valence sulfur in marine benthic habitats. *Environmental Microbiology* 16:3416-3430. 10.1111/1462-2920.12410

Pruesse E, Peplies J, and Glöckner FO. 2012. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28:1823-1829. 10.1093/bioinformatics/bts252

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, and Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 41:D590-D596. 10.1093/nar/gks1219

R Core Team. 2015. R: A Language and Environment for Statistical Computing. *Available at* http://www.R-project.org/.

Ravenschlag K, Sahm K, Pernthaler J, and Amann R. 1999. High bacterial diversity in permanently cold marine sediments. *Applied and Environmental Microbiology* 65:3982-3989.

Ruff SE, Biddle JF, Teske AP, Knittel K, Boetius A, and Ramette A. 2015. Global dispersion and local diversification of the methane seep microbiome. *Proceedings of the National Academy of Sciences* 112:4015-4020.

Šafařík I, and Šafaříková M. 1999. Use of magnetic techniques for the isolation of cells. *Journal of Chromatography B* 722:33-53.

Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, and Walker AW. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 12:87. 10.1186/s12915-014-0087-z

Schattenhofer M, and Wendeberg A. 2011. Capturing Microbial Populations for Environmental Genomics. *Handbook of Molecular Microbial Ecology I*: John Wiley & Sons, Inc., 735-740.

Schreiber L, Holler T, Knittel K, Meyerdierks A, and Amann R. 2010. Identification of the dominant sulfate-reducing bacterial partner of anaerobic methanotrophs of the ANME-2 clade. *Environmental Microbiology* 12:2327-2340. 10.1111/j.1462-2920.2010.02275.x

Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, and Huttenhower C. 2011. Metagenomic biomarker discovery and explanation. *Genome Biol* 12:R60. 10.1186/gb-2011-12-6-r60

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, and Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498-2504. 10.1101/gr.1239303

Sorokin DY, Kovaleva OL, Tourova TP, and Muyzer G. 2010. Thiohalobacter thiocyanaticus gen. nov., sp. nov., a moderately halophilic, sulfur-oxidizing gammaproteobacterium from hypersaline lakes, that utilizes thiocyanate. *Int J Syst Evol Microbiol* 60:444-450. 10.1099/ijs.0.012880-0

Steele JA, Countway PD, Xia L, Vigil PD, Beman JM, Kim DY, Chow C-ET, Sachdeva R, Jones AC, Schwalbach MS, Rose JM, Hewson I, Patel A, Sun F, Caron DA, and Fuhrman JA. 2011. Marine bacterial, archaeal and protistan association networks reveal ecological linkages. *ISME J* 5:1414-1425.

Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G, Djahanschiri B, Zeller G, Mende DR, Alberti A, Cornejo-Castillo FM, Costea PI, Cruaud C, d'Ovidio F, Engelen S, Ferrera I, Gasol JM, Guidi L, Hildebrand F, Kokoszka F, Lepoivre C, Lima-Mendez G, Poulain J, Poulos BT, Royo-Llonch M, Sarmento H, Vieira-Silva S, Dimier C, Picheral M, Searson S, Kandels-Lewis S, coordinators TO, Bowler C, de Vargas C, Gorsky G, Grimsley N, Hingamp P, Iudicone D, Jaillon O, Not F, Ogata H, Pesant S, Speich S, Stemmann L, Sullivan MB, Weissenbach J, Wincker P, Karsenti E, Raes J, Acinas SG, and Bork P. 2015. Structure and function of the global ocean microbiome. *Science* 348. 10.1126/science.1261359

Takano Y, Kaneko M, Kahnt Jr, Imachi H, Shima S, and Ohkouchi N. 2013. Detection of coenzyme F430 in deep sea sediments: A key molecule for biological methanogenesis. *Organic Geochemistry* 58:137-140. http://dx.doi.org/10.1016/j.orggeochem.2013.01.012

Trembath-Reichert E, Green-Saxena A, and Orphan VJ. 2013. Chapter Two - Whole Cell Immunomagnetic Enrichment of Environmental Microbial Consortia Using rRNA-Targeted Magneto-FISH. In: Edward FD, ed. *Microbial Metagenomics, Metatranscriptomics, and Metaproteomics*. San Diego, CA: Academic Press, 21-44.

Treude T, Orphan V, Knittel K, Gieseke A, House CH, and Boetius A. 2007. Consumption of Methane and CO2 by Methanotrophic Microbial Mats from Gas Seeps of the Anoxic Black Sea. *Applied and Environmental Microbiology* 73:2271-2283. 10.1128/aem.02685-06

Vigneron A, L'Haridon Sp, Godfroy A, Roussel EG, Cragg BA, Parkes RJ, and Toffin L. 2015. Evidence of Active Methanogen Communities in Shallow Sediments of the Sonora Margin Cold Seeps. *Applied and Environmental Microbiology* 81:3451-3459. 10.1128/aem.00147-15

von Wintzingerode F, Selent B, Hegemann W, and Göbel UB. 1999. Phylogenetic Analysis of an Anaerobic, Trichlorobenzene-Transforming Microbial Consortium. *Applied and Environmental Microbiology* 65:283-286.

Wagner M, Roger AJ, Flax JL, Brusseau GA, and Stahl DA. 1998. Phylogeny of Dissimilatory Sulfite Reductases Supports an Early Origin of Sulfate Respiration. *Journal of Bacteriology* 180:2975-2982.

Wegener G, Bausch M, Holler T, Thang NM, Prieto Mollar X, Kellermann MY, Hinrichs K-U, and Boetius A. 2012. Assessing sub-seafloor microbial activity by combined stable isotope probing with deuterated water and 13C-bicarbonate. *Environmental Microbiology* 14:1517-1527. 10.1111/j.1462-2920.2012.02739.x

Wegener G, Krukenberg V, Riedel D, Tegetmeyer HE, and Boetius A. 2015. Intercellular wiring enables electron transfer between methanotrophic archaea and bacteria. *Nature* 526:587-590. 10.1038/nature15733

White JR, Nagarajan N, and Pop M. 2009. Statistical methods for detecting differentially abundant features in clinical metagenomic samples. *PLoS Comput Biol* 5:e1000352. 10.1371/journal.pcbi.1000352

Wright JJ, Mewis K, Hanson NW, Konwar KM, Maas KR, and Hallam SJ. 2014. Genomic properties of Marine Group A bacteria indicate a role in the marine sulfur cycle. *ISME J* 8:455-468. 10.1038/ismej.2013.152

Yanagawa K, Sunamura M, Lever MA, Morono Y, Hiruta A, Ishizaki O, Matsumoto R, Urabe T, and Inagaki F. 2011. Niche separation of methanotrophic archaea (ANME-1 and-2) in methane-seep sediments of the eastern Japan Sea offshore Joetsu. *Geomicrobiology Journal* 28:118-129.

Yilmaz S, Haroon MF, Rabkin BA, Tyson GW, and Hugenholtz P. 2010. Fixation-free fluorescence in situ hybridization for targeted enrichment of microbial populations. *ISME J* 4:1352-1356.

Supplemental Table 1: Extracted DNA concentration per sample measured by fluorometer.

| Sample | Extracted DNA (ng/µl) |
|---|---|
| eel932BC1 | BD |
| eel932BC2 | BD |
| eel932BC3 | BD |
| Seep1a1441BC1 | BD |
| Seep1a1441BC2 | BD |
| DSS658BC1 | BD |
| DSS658BC2 | BD |
| DSS658BC3 | BD |
| Delta495aBC1 | BD |
| Delta495aBC2 | 0.05 |
| FixedBulk1 | 0.52 |
| FixedBulk2 | 0.55 |
| FixedBulk3 | 0.70 |
| UnfixedBulk2 | 0.80 |
| UnfixedBulk3 | 1.34 |

Supplemental Table 2: Expected and recovered sequence abundances among the mock communities show differential taxonomic biases. *Fold Change* is calculated by dividing the experimentally recovered relative abundance by the expected relative abundance. Four mock communities were designed with a selection of common methane seep bacterial and archaeal taxa at different relative abundance ratios. Mock community analysis revealed that relative abundances of *Helicobacteraceae* (*Sulfurovum*), *Desulfobacteraceae* (Seep-SRB1) and *Desulfobulbaceae* (*Desulfobulbus*) had little amplification bias as compared to other mock community taxa (fold change ranges 0.93-1.42, where 1.00 means expected relative abundance was returned). ANME-1b plasmids were also overall well represented (fold change 0.64 to 1.42) across the range of expected relative abundances (1% to 20%). In contrast, ANME-2a/b and ANME-2c plasmids were always under amplified in all of the mock communities (fold change 0.32 to 0.81). These results do not appear to correlate to primer hits in the SILVA SSU r123 database, where 89.5% of ANME-2c sequences were hit by 515f and 87.1% by 806r, but 94.3% of ANME-2a/b were hit by 515f and 806r. ANME-2a/b was a better match to the EMP primers, but both taxa were under amplified in mock community analysis. Amplification bias was not always uniform, where some templates saw varied amplification response depending on initial relative abundance in the mock community. The ANME-1a plasmid was over-amplified (3.35-2.44 fold change; Sup Table 2) when the plasmid was at 5% relative abundance and lower (Mock Communities 1-3). However, Mock Community 4 with the highest relative abundance (20%) of ANME-1a plasmids saw templates amplified to the expected relative abundance (0.97 fold change). *Thaumarchaeota*: Miscellaneous Crenarchaeota Group followed a similar pattern to ANME-1a: where it was 1% expected relative abundance, the fold change is ~5, and where it was 10% expected relative abundance, the fold change was less pronounced (~1.5). MBG-D sequences were slightly over amplified when at 1% expected relative abundance, and slightly under amplified when at 42% relative abundance. Bias was consistent across mock community samples when the relative percentage of that group (e.g. *Thermoplasmatales*, 40%) was the same in both samples. This suggests that analysis based on relative abundance between samples can be applied as a means of comparison, as long as the environmental OTUs of interest are above the detection threshold. A study of EMP primers with a pelagic marine community also reported discrepancies between mock community bias and independently assessed environmental sample bias for a dominant community members (Parada et al. 2015). Parada et al. similarly conclude that over-amplification of certain community members, in their case *Gammaproteobacteria*, was the cause of lower

than expected recovery, rather than lack of SAR11 and SAR116. Our ANME-2c results, therefore, serve as yet another example of how key community members can be under-represented when exploring unknown microbial systems. The severity of this issue for future studies is dependent on the research question, interpretation approach, and the phylogenic bias imparted on community members of interest. The phenomenon of less pronounced bias when templates are at higher starting relative abundances could be explained by the reannealing inhibition affect of high copy number templates in mock samples (Suzuki & Giovannoni 1996). Due to low template of Magneto-FISH samples, PCRs were done for a total of 35 cycles. Since bias is positively correlated with number of cycles (Suzuki & Giovannoni 1996), lowering PCR amplification cycles could improve bias issues. The lack of statistically significant ANME-2c correlations is expected since this group was recovered in so few samples. ANME-1a, however, may suffer from the opposite problem where over-amplification in iTag datasets reduces the ability to determine patterns with other OTUs. As an analogy, if the ANME-2c population is an image with only a few pixels and the image of the ANME-1a population is an image with oversaturated pixels, then neither has a workable dynamic range for correlation analysis. The log transform operation performed on the data before correlation analysis can reduce the bias between high and low abundance OTUs to some degree, but may not be sufficient in all cases, such as with these two OTUs. Several approaches can ameliorate some of the issues within iTag sequencing datasets: (1) Optimization of PCR conditions and use of high-fidelity DNA polymerase for amplification in conjunction with the (2) creation and sequencing of a mock community, if there is *a priori* knowledge of the community composition; (3) Data transformation(s) before statistical analysis (i.e., square root, fourth root, or log transformations) and (4) examining the behavior of single OTUs across multiple samples/treatments may be more robust than direct comparison of OTUs within a single sample; (5) Whole-community comparisons (i.e. UniFrac, ANOVA, ANOSIM) to minimize single-taxon biases by including all taxa.

| Plasmid Taxonomy | | Mock 1 | Mock 2 | Mock 3 | Mock 4 |
|---|---|---|---|---|---|
| ***Desulfobulbaceae* (DSB)** | *Expected* | ***3.0%*** | ***3.0%*** | ***11.0%*** | ***11.0%*** |
| | Experimental | 3.0% | 3.2% | 11.6% | 12.9% |
| | **Fold Change** | **0.99** | **1.05** | **1.05** | **1.17** |
| | Std. Dev. | 0.06 | | 0.10 | |
| ***Helicobacteraceae*** | *Expected* | ***25.0%*** | ***25.0%*** | ***1.0%*** | ***1.0%*** |
| | Experimental | 28.0% | 23.4% | 1.4% | 1.2% |
| | **Fold Change** | **1.12** | **0.93** | **1.38** | **1.17** |
| | Std. Dev. | 0.18 | | 0.16 | |
| ***Desulfobacteraceae* (DSS)** | *Expected* | ***9.0%*** | ***9.0%*** | ***31.0%*** | ***31.0%*** |
| | Experimental | 12.7% | 11.0% | 34.8% | 31.5% |
| | **Fold Change** | **1.42** | **1.22** | **1.12** | **1.02** |
| | Std. Dev. | 0.15 | | 0.10 | |
| **ANME-1a** | *Expected* | *1.0%* | *4.0%* | *5.0%* | *20.0%* |
| | Experimental | 3.4% | 9.9% | 14.5% | 19.4% |
| | **Fold Change** | **3.35** | **2.48** | **2.90** | **0.97** |
| **ANME-1b** | *Expected* | ***1.0%*** | ***4.0%*** | ***5.0%*** | ***20.0%*** |
| | Experimental | 1.4% | 3.4% | 6.9% | 12.8% |
| | **Fold Change** | **1.42** | **0.85** | **1.37** | **0.64** |

| | | | | | |
|---|---|---|---|---|---|
| **ANME-2a/b** | *Expected* | ***6.0%*** | ***2.0%*** | ***30.0%*** | ***10.0%*** |
| | Experimental | 3.7% | 0.6% | 10.2% | 6.2% |
| | **Fold Change** | **0.61** | **0.32** | **0.34** | **0.62** |
| **ANME-2c** | *Expected* | ***3.5%*** | ***1.0%*** | ***15.0%*** | ***5.0%*** |
| | Experimental | 1.9% | 0.5% | 6.9% | 4.1% |
| | **Fold Change** | **0.55** | **0.49** | **0.46** | **0.81** |
| **Miscellaneous Crenarchaeota Group** | *Expected* | ***10.0%*** | ***10.0%*** | ***1.0%*** | ***1.0%*** |
| | Experimental | 14.8% | 14.9% | 5.9% | 4.7% |
| | **Fold Change** | **1.48** | **1.49** | **5.87** | **4.69** |
| | Std. Dev. | 0.01 | | 0.22 | |
| ***Thermoplasmatales*** | *Expected* | ***41.5%*** | ***42.0%*** | ***1.0%*** | ***1.0%*** |
| | Experimental | 30.8% | 30.7% | 6.5% | 6.1% |
| | **Fold Change** | **0.74** | **0.73** | **6.48** | **6.09** |
| | Std. Dev. | 0.01 | | 0.06 | |

Supplemental Table 3: Sequences per sample post processing. Total sequences per sample after mothur processing and 0.1% bulk sediment cutoff and total sequences remaining for the most abundant ANME-2c, SEEP-SRB1, and ANME-1a OTUs. We also performed a BLASTN ( Madden 2002 ) search of all contigs from all samples against an in-house database of 155 ANME-2c 16S rRNA sequences of >500 bp. This yielded 1,395 iTag sequences with an e-value greater than or equal to $10^{-130}$, corresponding to 99-100% sequence identity match to sequences from our ANME-2c database. We then tracked this set of BLAST match contigs through each step in the mothur pipeline, with a final result of 1,260 sequences remaining in this BLAST set from the original contig file. Thus 92% of our ANME-2c BLAST hit set remained through the mothur processing pipeline. This suggests that the lack of ANME-2c sequences in our downstream database was not due to spurious removal during sequence processing.

| | after 0.1% removal | ANME-2c | SEEP-SRB1 | ANME-1a |
|---|---|---|---|---|
| Sample ID | Sequences | Otu16818 | Otu17765 | Otu12964 |
| ETR-D1-DSS658BC1 | 3460 | 0 | 120 | 67 |
| ETR-D2-DSS658BC2 | 5310 | 0 | 193 | 78 |
| ETR-D3-DSS658BC3 | 9859 | 0 | 393 | 804 |
| ETR-E1-eel932BC1 | 6304 | 0 | 279 | 36 |
| ETR-E2-eel932BC2 | 5253 | 0 | 100 | 176 |
| ETR-E3-eel932BC3 | 3293 | 0 | 414 | 52 |
| ETR-F1-FixedBulk1 | 12257 | 21 | 171 | 1263 |
| ETR-F2-FixedBulk2 | 4957 | 63 | 70 | 596 |
| ETR-F3-FixedBulk3 | 14400 | 27 | 215 | 1192 |
| ETR-S1-Seep1a1441BC1 | 2562 | 0 | 200 | 64 |
| ETR-S2-Seep1a1441BC2 | 7867 | 0 | 408 | 865 |
| ETR-U2-UnfixedBulk2 | 14441 | 54 | 309 | 1124 |
| ETR-U3-UnfixedBulk3 | 11447 | 32 | 244 | 1180 |
| ETR-delta1-Delta495aBC1 | 1552 | 0 | 0 | 0 |
| ETR-delta2-Delta495aBC2 | 3790 | 0 | 91 | 0 |
| Total | 171154 | 197 | 3207 | 7497 |
| Percent of Unfixed Bulk Sediment | - | 0.33% | 2.14% | 8.90% |

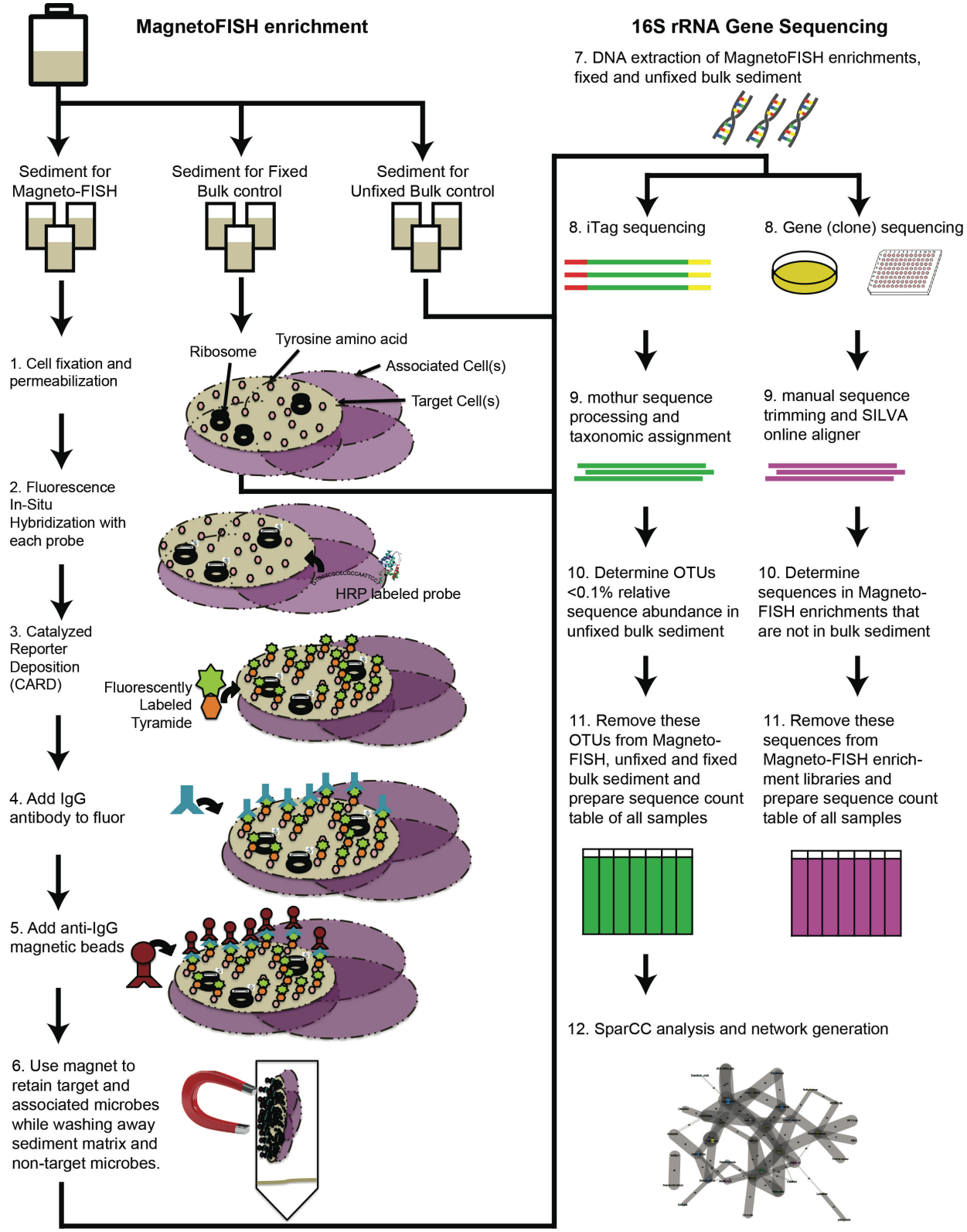
Supplemental Table 4: Mock Community sequencing error rates (0.025-0.095%; Sup Table 4) were of the same magnitude as Kozich et al. (~0.01%, 2013). Rarefaction of the mock community to 5,000 sequences shows OTU inflation rates of 3 to 4 times expected number of OTUs, after 97% OTU clustering and removal of singletons. The inflation rate is calculated by total number of OTUs recovered divided by original number of template plasmids. Since our environmental mock community only had 12 templates, the number of spurious OTUs is expected to be high (Huse et al. 2010). Experimental sediment samples have 10 to 100 times more templates, so inflation rates are expected to be much lower (10-1%).

| Sample | OTU@5000 | OTU@5000, singletons remv. | OTU@5000 singletons remv. inflation rate | Error Rate |
|---|---|---|---|---|
| Mock 1 | 65 (57-73) | 32 (28-36) | 2.7 | 0.025% |
| Mock 2 | 58 (50-68) | 30 (26-34) | 2.5 | 0.033% |
| Mock 3 | 113 (102-123) | 38 (35-40) | 3.2 | 0.095% |
| Mock 4 | 85 (77-94) | 38 (35-41) | 3.2 | 0.085% |

Supplemental Table 5: All associations that occurred in at least 50 out of 100 networks for combined Magneto-FISH and bulk sediment samples with their number of occurrences, average correlation, and p-values.

| OTU_1 | Tax_1 | OTU_2 | Tax_2 | Occurrence | Correlation | Pvalue |
|---|---|---|---|---|---|---|
| Otu072 | ANME-1b | Otu090 | JS1 | 100 | 0.746 | 0.002 |
| Otu069 | OD1 | Otu077 | Pelobacter | 100 | 0.872 | 0 |
| Otu014 | AKAU3564_sed. | Otu082 | SEEP-SRB1 | 100 | 0.729 | 0.001 |
| Otu071 | Hyd24-12 | Otu083 | SEEP-SRB1 | 100 | 0.765 | 0 |
| Otu073 | Sulfurovum | Otu087 | Sulfurovum | 100 | 0.798 | 0.001 |
| Otu134 | MBG-B | Otu135 | OM1 | 99 | 0.737 | 0.001 |
| Otu123 | VC2.1_Bac22 | Otu128 | WS3 | 98 | 0.718 | 0.001 |
| Otu106 | OD1 | Otu115 | Desulfocapsa | 97 | 0.705 | 0.002 |
| Otu071 | Hyd24-12 | Otu097 | Spirochaeta | 96 | 0.708 | 0.002 |
| Otu083 | SEEP-SRB1 | Otu097 | Spirochaeta | 96 | 0.691 | 0.001 |
| Otu100 | OD1 | Otu107 | Anaerolineaceae | 95 | 0.709 | 0.001 |
| Otu021 | WF1B-44 | Otu022 | Anaerolineaceae | 95 | 0.681 | 0.001 |
| Otu068 | Desulfoluna | Otu119 | VC2.1_Bac22 | 95 | 0.692 | 0.001 |
| Otu077 | Pelobacter | Otu094 | BD2-2 | 94 | 0.703 | 0.002 |
| Otu073 | Sulfurovum | Otu101 | SEEP-SRB1 | 94 | 0.778 | 0.003 |
| Otu097 | Spirochaeta | Otu131 | Actibacter | 93 | -0.688 | 0.002 |
| Otu014 | AKAU3564_sed. | Otu089 | SEEP-SRB1 | 93 | 0.691 | 0.002 |
| Otu122 | SEEP-SRB4 | Otu131 | Actibacter | 92 | 0.679 | 0.002 |
| Otu112 | SEEP-SRB1 | Otu114 | JS1 | 92 | 0.663 | 0.003 |
| Otu023 | Anacalomicrobium | Otu024 | MSBL8 | 92 | 0.684 | 0.001 |
| Otu102 | SEEP-SRB1 | Otu116 | VC2.1_Bac22 | 92 | 0.701 | 0.001 |
| Otu098 | Desulfobulbus | Otu108 | BD2-2 | 91 | -0.668 | 0.003 |
| Otu069 | OD1 | Otu086 | OM1 | 91 | 0.688 | 0.003 |
| Otu090 | JS1 | Otu114 | JS1 | 90 | 0.685 | 0.002 |
| Otu095 | SEEP-SRB2 | Otu119 | VC2.1_Bac22 | 90 | 0.687 | 0.001 |
| Otu069 | OD1 | Otu094 | BD2-2 | 89 | 0.683 | 0.002 |
| Otu091 | SEEP-SRB2 | Otu114 | JS1 | 89 | 0.664 | 0.004 |
| Otu082 | SEEP-SRB1 | Otu089 | SEEP-SRB1 | 89 | 0.69 | 0.002 |
| Otu071 | Hyd24-12 | Otu100 | OD1 | 86 | 0.679 | 0.003 |
| Otu091 | SEEP-SRB2 | Otu107 | Anaerolineaceae | 85 | 0.688 | 0.003 |
| Otu087 | Sulfurovum | Otu099 | Sulfurovum | 85 | 0.675 | 0.004 |
| Otu094 | BD2-2 | Otu100 | MBG-B | 84 | 0.662 | 0.003 |
| Otu081 | Caldthrix | Otu086 | OM1 | 84 | -0.642 | 0.003 |
| Otu127 | Anaerolineaceae | Otu132 | Thiohalobacter | 84 | 0.651 | 0.002 |
| Otu098 | Desulfobulbus | Otu122 | SEEP-SRB4 | 81 | 0.66 | 0.003 |
| Otu091 | SEEP-SRB2 | Otu134 | MBG-B | 80 | 0.663 | 0.003 |
| Otu105 | CS-B046 | Otu108 | BD2-2 | 79 | 0.656 | 0.003 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Otu077 | Pelobacter | Otu086 | OM1 | 79 | 0.655 | 0.003 |
| Otu073 | Sulfurovum | Otu098 | Desulfobulbus | 78 | 0.677 | 0.004 |
| Otu016 | Anaerolineaceae | Otu092 | SB-5 | 77 | 0.649 | 0.003 |
| Otu068 | Desulfoluna | Otu095 | SEEP-SRB2 | 77 | 0.658 | 0.002 |
| Otu103 | ANME-2a/b | Otu135 | OM1 | 71 | 0.647 | 0.004 |
| Otu126 | WS3 | Otu129 | Clostridia | 70 | 0.653 | 0.003 |
| Otu090 | JS1 | Otu100 | OD1 | 69 | 0.647 | 0.004 |
| Otu122 | SEEP-SRB4 | Otu128 | WS3 | 68 | 0.641 | 0.004 |
| Otu027 | Anaerolineaceae | Otu030 | WCHB1-69 | 66 | 0.657 | 0.001 |
| Otu112 | SEEP-SRB1 | Otu121 | Desulfarc._unclt. | 65 | -0.648 | 0.004 |
| Otu084 | Desulfobulbus | Otu120 | Desulfococcus | 65 | 0.663 | 0.006 |
| Otu051 | Lutibacter | Otu096 | ANME-1b | 64 | -0.636 | 0.004 |
| Otu005 | Desulfocapsa | Otu032 | Hyd24-12 | 64 | 0.648 | 0.003 |
| Otu011 | Anaerolineaceae | Otu072 | ANME-1b | 63 | -0.639 | 0.003 |
| Otu072 | ANME-1b | Otu100 | OD1 | 59 | 0.637 | 0.005 |
| Otu072 | ANME-1b | Otu094 | BD2-2 | 58 | 0.646 | 0.003 |
| Otu124 | WS3 | Otu134 | MBG-B | 57 | 0.631 | 0.003 |
| Otu020 | pMC2A209 | Otu051 | Lutibacter | 55 | 0.646 | 0.003 |
| Otu113 | Desulfobulbus | Otu118 | SAR406 | 55 | 0.643 | 0.003 |
| Otu036 | Desulfocapsa | Otu046 | Sulfurimonas | 54 | 0.646 | 0 |
| Otu069 | OD1 | Otu081 | Caldthrix | 52 | -0.633 | 0.004 |
| Otu108 | BD2-2 | Otu120 | Desulfococcus | 50 | 0.632 | 0.003 |
| Otu103 | ANME-2a/b | Otu122 | SEEP-SRB4 | 50 | 0.635 | 0.004 |
| Otu046 | Sulfurimonas | Otu054 | WCHB1-69 | 50 | 0.636 | 0.001 |




Supplemental Figure 1: Methods flow diagram.


| Functional Gene | dsrA | aprA | soxB |
|---|---|---|---|
| Desulfobacteraceae (DSB) | 15 | 7 | |
| Desulfobulbaceae (DBB) | 1 | 0 | |
| Gammproteobacteria (Endosymbiont) | | 5 | |
| Other (Cluster B) | | 1 | |
| Epsilonproteobacteria (Sulfurimonas) | | | 20 |
| Epsilonproteobacteria (Sulfurovum) | | | 2 |

Supplemental Figure 2: Phylogenetic trees of SoxB, AprA, and DsrA functional genes from a MSMX-Eel_932 Magneto-FISH enrichment. aLRT SH-like values above 50% displayed for branch support. Similar sequence clusters represented by one sequence are indicated in parentheticals, and sequences from this study are in bold. Clones recovered are also summarized in table form. As another method of assessing Magneto-FISH diversity, we examined functional genes relating to the sulfur cycle. This method can also provide insight into phylogenetic connections between 16S rRNA and sulfur cycling functional genes. Clone libraries were constructed from an MSMX-Eel_932 Magneto-FISH capture and performed with sediment as iTag libraries (see Materials & Methods). The following genes relating to sulfur cycling pathways were chosen for this analysis: soxB (sulfur oxidation, protein-S-thiocysteine sulfate hydrolase), aprA (sulfur oxidation and reduction, adenylylsulfate reductase α subunit), and dsrA (sulfur oxidation and reduction, dissimilatory sulfite reductase). Phylogenetic analysis of soxB clones from the MSMX-Eel_932 Magneto-FISH returned only *Epsilonproteobacteria*sequences from both *Sulfurovum* (2 clones) and *Sulfurimonas* (20 clones) clades (Sup Figure 1). From a total of 13 aprA clones, 7 were retrieved from the *Desulfobacteraceae* clade (SEEP-SRB1 containing), none from the *Desulfobulbaceae*, 5 from *Gammaproteobacteria* Endosymbiont clade, and 1 from the "Cluster B" GoM clone clade ( Meyer & Kuever 2007 ) . 15 of 16 dsrA clones were from the *Desulfobacteraceae* clade, with one clone from the *Desulfobulbaceae* clade ( Müller et al. 2015 ) . Functional gene clone libraries were not only successful in providing another means to assess Magneto-FISH enrichment, but provide an example of how this technique can be utilized to target specific 16S rRNA populations and the metabolic diversity contained. This is particularly useful in cases where 16S rRNA and functional gene phylogenies are not well aligned.